

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 23:09:34 ; Search time 6466 Seconds

(without alignments)
10723.547 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1482
Sequence: 1 CAGACACACAAAGCGATCT.....AATATTTCGAACATATA 1482Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hc:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_ges1:*

10: gb_ges2:*

11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	291.4	19.7	317	10	T09652 0224m3 gmbp
2	256	17.3	1721	8	DR145459 49056454
3	249.4	16.8	1808	8	DR125630 49096520
4	248.4	16.8	1519	10	AG386893 Mus muscu
5	247.2	16.7	1453	10	AJ591978 Arabidops
6	244.4	16.5	2071	8	DR147304 49027161
7	243	16.4	1932	8	DR141720 49300102
8	241	16.3	1727	10	CL114085 ISB1-59J1
9	240	16.2	1491	10	CG753221 P048-2-A0
10	238.4	16.1	1783	8	DR141824 49190649
11	237.8	16.0	1434	10	AJ592058 Arabidops
12	237.8	16.0	1808	10	CL078963 CH216-154
13	235	15.9	1654	10	CL118709 ISB1-7211
14	234.2	15.8	1797	10	CL066150 CH216-108
15	234	15.8	1829	10	CL109409 ISB1-51H2
16	233.8	15.8	1736	10	CL066373 CH216-108
17	233.6	15.8	1407	10	AJ592026 Arabidops
18	233.6	15.8	1753	8	DR131208 49217594
19	233.8	15.7	1512	10	CL113943 ISB1-59E2
20	232.8	15.7	1531	10	CG748014 P041-4-B0
21	232.6	15.7	1811	10	CG753732 P048-4-G0
22	231.6	15.6	1917	10	CL081899 CH216-164

23	231.4	15.6	1650	10	CL041163	CL041163 CH216-51P
24	231.2	15.6	1762	10	CL082587	CL082587 CH216-168
25	230.4	15.5	1824	10	CL081234	CL081234 CH216-161
26	230.2	15.5	1727	10	CL118815	CL118815 ISB1-7205
27	230	15.5	1632	10	CL082569	CL082569 CH216-167
28	229.8	15.5	1616	10	CL081995	CL081995 CH216-165
29	229.6	15.5	1630	10	CL078686	CL078686 CH216-152
30	229.2	15.5	1457	10	CL082658	CL082658 CH216-169
31	228.8	15.4	1566	10	CG757757	CG757757 P053-1-D0
32	228.6	15.4	1434	9	CC187638	CC187638 CH261-98P
33	228.4	15.4	1507	10	AG346189	AG346189 Mus muscu
34	228.4	15.4	1823	10	CL058635	CL058635 CH216-88H
35	228	15.4	1259	11	CR004722	CR004722 Forward s
36	228	15.4	2033	10	CL081977	CL081977 CH216-165
37	228	15.4	2103	10	CL041093	CL041093 CH216-51M
38	227.6	15.4	1712	10	CL078539	CL078539 CH216-151
39	226.8	15.3	1547	10	CL066420	CL066420 CH216-108
40	226.6	15.3	1377	10	CL077307	CL077307 CH216-144
41	226.4	15.3	1716	9	CC222065	CC222065 CH261-11A
42	225.6	15.2	1562	10	CL078533	CL078533 CH216-151
43	225.6	15.2	1651	10	CL079030	CL079030 CH216-154
44	225.4	15.2	1482	10	AG332292	AG332292 Mus muscu
45	224.2	15.1	1506	10	CL082761	CL082761 CH216-171

ALIGNMENTS

RESULT 1
T09652 317 bp DNA linear GSS 05-DEC-2000
LOCUS 0224m3 gmbpFHB3.1, G. Roman Reddy Plasmodium falciparum genomic
DEFINITION clone 0224m, genomic survey sequence.
ACCESSION T09652
VERSION T09652.1 GI:319484
KEYWORDS GSS.

ORGANISM Plasmodium falciparum (malaria parasite P. falciparum)

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 317)
Reddy,G.R., Chakrabarti,D., Schuster,S.M., Ferl,R.J., Almira,E.C.
and Dame,J.B.TITLE Gene sequence tags from Plasmodium falciparum genomic DNA fragments
prepared by the genease activity of mung bean nucleaseJOURNAL Proc. Natl. Acad. Sci. U.S.A. 90, 9867-9871 (1993)
PUBMED 8234327COMMENT Other GSSs: 0224m7
Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.ufl.edu
Seg primer: T3
Class: Shotgun.

FEATURES

source

1. 317
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/clone="0224m"
/lab_host="E. coli XL1-Blue"
/clone_id="gmbpFHB3.1, G. Roman Reddy"
/note="Vector: pBluescript SK(+); Genomic DNA, from
asynchronous blood stage parasites of the cloned Honduran
HB3 isolate cultured in vitro, was digested at 500C
(Vernick, K.D., Imberst, R.B., and McCutchan, T.F. 1988.
Nucleic Acids Research 16:6883-6896). The ends of the
fragments were polished using T4 DNA polymerase, and the
fragments were ligated to EcoR V-cleaved and
dephosphorylated pBluescript SK(+). Recombinant plasmids

RESULT 3	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	COMMENT
Oy	902	GACATCTTGAAGAAAGAAAGATGGTTCATATAAACCGAACAAGAAAGACATTAACCTG	1096	1037	AA								
Oy	962	CTGACATCAAAATCATCATTTAGAGACAGTAATATTTCTGATGTTATGATTTTCAAA	1021	1097	AA								
Oy	1022	TAACTAAGTATGAGATGATAATTAAGTCTGAATATGACGATTCATTAATAGATAGAG	1081	1157	AA								
Oy	1082	AAGATGATGAAGACTTACAGCAATTTAGCCTATTGTCGACATTAAGCAATTTCCAGATG	1141	1217	AA								
Oy	1142	AAGAAACATGATGAATTTATTAAGAACTGAGAAGTTTGATGAGAAAAATGAAATTTAG	1201	1277	AA								
Oy	1302	ATGATTTAGATGAAGAAATAGAAAAATCATCAGAAATTAATCTGAAGAAAAATTAATA	1261	1337	AA								
Oy	1262	AAGCAAGAAATTTGAAAAAACAAAGCTAATTAATTTTAAACCAATGATTAAGTTGT	1321	1397	AA								
Oy	1322	ATGATGACATATTTAAAAATATATTAATGATTAAGCAGTTTAATAGCAAAAGAAAAAT	1381	1457	AA								
Oy	1382	TCATAAAA 1389		1517	AAAAAAAA 1524								
RESULT 3	DR125630	1808 bp mRNA linear EST 15-JUN-2005											
LOCUS	49096520	Drosophila pseudoobscura embryonic cDNA library Drosophila											
DEFINITION	pseudobscura cDNA clone 023 5', mRNA sequence.												
ACCESSION	DR125630												
VERSION	DR125630.1	GI:67842328											
KEYWORDS	EST.												
SOURCE	Drosophila pseudoobscura												
ORGANISM	Drosophila pseudoobscura												
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;												
AUTHORS	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;												
	Ephydroidea; Drosophilidae; Drosophila.												
	1 (bases 1 to 1808)												
	Richards S., Liu Y., Beltencourt B.R., Hradecky P., Letovsky S.,												
	Nielsen R., Thornton K., Hubisz M.J., Chen R., Meisel R.P.,												
	Conrione O., Hua S., Smith M.A., Zhang P., Liu J., Bussemaker H.J.,												
	van Batenburg M.F., Howells S.L., Scherer S.E., Sodergren B.,												
	Matthews B.B., Crosby M.A., Schroeder A.J., Ortiz-Barrileto D.,												
	Rives C.M., Metzger M.L., Muzny D.M., Scott G., Steffen D.,												
	Wheeler D.A., Worley K.C., Hayak P., Durbin K.J., Bgan A.,												
	Gill R., Hume J., Morgan M.B., Miner G., Hamilton C., Huang Y.,												
	Waldron L., Verdusco D., Cleic-Blankenburg K.P., Dubchak I.,												
	Noor M.A., Anderson W., White K.P., Clark A.G., Schaeffer S.W.,												
	Galbraith W., Weinstein G.M. and Gibbs R.A.												
	Comparative genome sequencing of Drosophila pseudoobscura:												
	Chromosome1, gene, and cis-element evolution												
	Genome Res. 15 (1), 1-18 (2005)												
	15632085												
	Contact: Stephen Richards												
	Human Genome Sequencing Center												
	Baylor College of Medicine												

FEATURES		Insert Length: 1750	Std Error: 0.25.
SOURCE		Location/Qualifiers	
		1..1808	
		/organism="Drosophila pseudoobscura"	
		/mol_type="mRNA"	
		/db_xref="taxon:7237"	
		/clone="023"	
		/dev_stage="0-18h embryos"	
		/clone_id="Drosophila pseudoobscura embryonic cDNA library"	
		/note="Vector: pOT87; Site_1: EcoRI; Site_2: XhoI; oligo dt priming from poly A+ RNA, directionally cloned"	
ORIGIN			
Query Match	16.8%;	Score 249.4;	DB 8; Length 1808;
Best Local Similarity	48.6%;	Pred. No. 3.1e-31;	
Matches 670;	Conservative	0; Mismatches 709;	Indels 0; Gaps 0;
QY	2	AAAGAACCAACGCGATCTAGAACAGAGAGCGTGTAAAGAAAAGTTGCAGAACAC	61
DB	340	AAA	399
QY	62	AAAGCGATTTAGACAGATGACTTGCTTAAGAAAAGTTACAGAGCAGCAAGCGATT	121
DB	400	AAA	459
QY	122	TAGAACAGAGAGACTTGCTTAAGAAAAGTTGCAGAGACCAACGCGATTTAGACAG	181
DB	460	AAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	519
QY	182	AGACGCTGCTAAAGAAAAGTTGCAGAACCAACGCGATTTAGACAGAGAGAGTG	241
DB	520	AAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAAAAAAAAAGAAAAGAAAAGAGAA	579
QY	242	CTTAAGAAAAGTTGCAGAGAACCAACGCGATTTAGAACAGATGACTTGCTTAAGAA	301
DB	580	AAAAAAAAAGAAAAGAAA	639
QY	302	AGTTACAAGACGCAACGCGATTTAGACAGAGAGCGTGTAAAGAAAAGTTGCAAG	361
DB	640	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAGAAAAG	699
QY	362	AACACAAAGCGATTTGACACAGAGAGCGTGTAAAGAAAAGTTGCAGAACCAACAA	421
DB	700	AAAAAAAAAGAAAAGAAAAAAAAAGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	759
QY	422	GCGATTTAGACAGAGAGACTTGCTTAAGAAAAGTTGCAGAACCAACGCGATTTAG	481
DB	760	AAAAAAAAAGAAAAGAAA	819
QY	482	AACACAGAGAGCGTGTAAAGAAAAGTTGCAGAACCAACGCGATTTAGAACAGAGA	541
DB	820	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAGAAAAGAA	879
QY	542	GACGTGCTAAAGAAAAGTTGCAGACCAACAGCGATTGAAACAAGAGAGCGTGCTA	601
DB	880	AAAAAAAAAAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAAAAAAAAAGAAAAG	939
QY	602	AAAGAAAAGTTGCAGACGCAACGCGATTTAGAACAAAGGAGCGGTATCGAAAAAA	661
DB	940	AAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAAAAAAAAAGAAAAGAAAAGAA	999
QY	662	ATTAGAAAAGAAAAGAACATGAGATATTAGCAGAGATTTATATGTCGTTTAA	721
DB	1000	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAGAAAAGAA	1059
QY	722	AAATACCAAGCTATGAACCTTCATCAGAAAAATGAACGTGATATTATATACCATCAAT	781
DB	1060	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAGAAAAGAA	1119
QY	782	CTTCTTTAATCTCAGAGCAACAGAGAGAAATGTAGAGTTCCAGAGAAATATCTTAATAG	841
DB	1120	AAAGAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAGAAAAAAAAAGAAAAGAA	1179

Oy	842	AAAAAACAATATAGACAACTCTTTCAACAAATGTTGAAGGACGAAGGATATACATTAAG	901
Db	1180	AAAGAAAAAGAAAAAGAGAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	1239
Oy	902	GACATCTTGAAGAAAAAGAAAGATGGTTCAATAAAAACAAGAAAAAGAGATTAATCTG	961
Db	1240	AAAAAAAAAAAAAAAAAGAAATAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAA	1299
Oy	962	CTGACATACAAAATCATATACATTAGACAGTAAATATTTCTGATGTTAATGATTTTCAA	1021
Db	1300	AA	1359
Oy	1022	TAACTAGATATAGAGATGAAATTAAGTCTGAAATATGACGATTCATTAATAGATGAAG	1081
Db	1360	AAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1419
Oy	1082	AAGATGATGAAGACTTAGACGATTTAAGCCTATTTGCAATATGACATTTCCAGATG	1141
Db	1420	AAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1479
Oy	1132	AAGAAAAACATAGCAATTTATATAAGAACTAGAAAGATTGATAGAGAAAAATGAAATTTTG	1201
Db	1480	AAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1539
Oy	1202	ATGATTTGATGACAGAAATAGAAAAATCATCGAAGAAATTTCTCAAGAAAAATTAATA	1261
Db	1540	AAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1599
Oy	1262	AAGCAAAAGAAATATGAAAAAAACAAGATATATATTTTAAACCAATGATTAAGTTGT	1321
Db	1600	GGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA	1659
Oy	1322	ATGATGACATCTTTTAAAAATTTAAATGATTAACAGTTAATATAGAAAAAGAAAA	1380
Db	1660	AAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA	1718
RESULT 4	AG386893/c		
LOCUS	AG386893	1519 bp	DNA linear GSS 21-DEC-2004
DEFINITION	Mus musculus molossinus DNA, clone:MSM901-201B07.TU, genomic survey		
ACCESSION	AG386893		
VERSION	AG386893.1	GI:47998098	
KEYWORDS	GSS.		
SOURCE	Mus musculus molossinus (Japanese wild mouse)		
ORGANISM	Mus musculus molossinus		
REFERENCE	Mus musculus molossinus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1. Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriawaki, K. and Shiroishi, T.		
JOURNAL	Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end		
PUBLISHED	Genome Res. 14 (12), 2439-2447 (2004)		
REFERENCE	2 (baaes 1 to 1519)		
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-NOV-2003) Masahito Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan		
COMMENT	(E-mail:hattori@gsc.riken.jp, URL:http://hnp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)		
	Clones are derived from the mouse BAC library MSM901. For BAC library availability, please contact Kuniya Abe (abe@crc.riken.jp).		
	Teikubio Institute, Bio Resource Center,		
	The Institute of Physical and Chemical Research (RIKEN) 3-1-1		
	Koyada, Teikuba, 305-0074 Japan		

FEATURES		source	location/Qualifiers
Query Match	16.8%; Score 248.4; DB 10; Length 1519;		
Best Local Similarity	45.7%; Pred. No. 4.6e-31;		
Matches 641; Conservative	0; Mismatches 756; Indels 7; Gaps 1		
<p>phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp PRIMERS Sequencing : TJ LIBRARY Vector : PBACE3.6 R.Site 1 : Ecoli R.Site 2 : Ecoli.</p>			
<p>1. .1519 /organism="Mus musculus molossinus" /mol_type="genomic DNA" /sub_species="molossinus" /db_xref="taxon:57486" /clone="MSM01-201E07.TJ" /sex="male" /tissue_type="mixture of kidney and spleen" /clone_id="MSM01 Mouse Male BAC Library"</p>			
ORIGIN			
Query	21	AGACCAAGAGACGCTGCTTAAGAAACTTCCAGAGACCAAGCGATTGACACAGA	80
Db	1516	AA	145
Query	81	TAGACTTCTTAAGAAAGATTACAAAGACAGACCAAGCGATTGACACAGAGACTTGC	140
Db	1456	AA	139
Query	141	TAAAGAAAGTTGCAAGACCAACAAAGCGATCTAGACAGAGAGAGCTGCTAAAGAAA	200
Db	1396	AA	133
Query	201	GTTGCAGACACACAAAGCGATTGACACAGAGAGAGCTGCTAAGAAAGTTGCTCAGA	260
Db	1336	AA	127
Query	261	ACAACAAAGCGATTGACAAAGATGAGCTGTAAAGAAAAGTTACAAAGACCAAG	320
Db	1276	AA	121
Query	321	CGATTAGACAGAGAGAGCTGCTAAAGAAAGTTGCAAGACCAAGCGATTGAGA	380
Db	1217	-----AA	116
Query	381	ACAAGAGAGAGCTGCTAAAGAAAGTTGCAAGACCAAGCGATTGACACAGAG	440
Db	1163	AA	110
Query	441	ACTGCTAAGAAAAGTTGCAAGACCAAGCGATTGACAAAGAGAGAGAGCTGCTA	500
Db	1103	AA	104
Query	501	AGAAAGTTGCAAGACCAAGCGATTGACAAAGAGAGAGCTGCTAAAGAAAGTT	560
Db	1043	AA	984
Query	561	GCAGACACACAAAGCGATTGACAAAGAGAGAGCTGCTAAGAAAAGTTGCAAGAGCA	620
Db	983	AA	924
Query	621	GCAAGAGATTGACAAAGAGAGCGCTGATCGAAGAAAAGTTGAAAGAAAAAGGA	680
Db	923	AA	864
Query	681	ACATGAGATATATTAGAGAGAGATTATAGTCTTTAGAAATACAGATATGAACT	740
Db	863	AA	804
Query	741	TCCATCAGAAAAATGAACTGATATATTATACACATCACTCTTTACTCAGACAA	800

[illegible]

COMMENT	FEATURES	ORIGIN
Gaston Cremieux, 91057 Evry cedex, FRANCE PCR was performed on DNA from transformants of <i>Arabidopsis thaliana</i> plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbsgap.versailles.inra.fr/publiclines/ . This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (http://www.genoplante.com and http://genoplante-info.infobio-gen.fr).	Location/Qualifiers 1..1453 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /clone="598D01" /clone_1fb="Arabidopsis thaliana T-DNA insertion lines" /ecotype="Wassilewskija" 1..1453 /note="T-DNA flanking sequence right border"	
Query Match Best Local Similarity 44.0%; Pred. No. 7.2e-31; Matches 618; Conservative 0; Mismatches 785; Indels 0; Gaps 0;		
2 AAGAACAAACAAGCATCTAGAACACAGAGACGCTGTAAAGAAAGTTGCCAAGACAC 61		
1419 ANNAAA 1360		
62 AAAGCGATTGAAACAAGATGACTGCTTAAGAAAAATTGACAAGACGACGAACGATT 121		
1359 AA 1300		
122 TAGAACAAAGAGAGCTTGCTAAACAAAAGTTGCAAGAACAAAGCGATTGACACAG 181		
1299 AA 1240		
182 AGAGACGTCTAAAGAAAAGTTGCAGAGAACACAAAGCGATTGAAACAAGAGACGTG 241		
1239 AA 1180		
242 CTTAAAGAAAAGTTGCAGAGAACAAAGCGCTTTGAGAACAGATGACTGCTAAAGAA 301		
1179 AA 1120		
302 AGTTACAGAGACCAAGCGATTGTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAG 361		
1119 AA 1060		
362 AACACAAAGCGATTGTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGACACAA 421		
1059 AA 1000		
422 GCGATTGTAGACAGAGAGCTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTAG 481		
999 AA 940		
482 AACACAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAG 541		
999 AA 880		
542 GACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGCGTCT 601		
879 AA 820		
602 AAGAAAGTTGCAAGAGACGAAAGAGATTGAGAACAAAGAGCGTGTATCGAAAAAA 661		
819 AA 760		


```

Qy 602 AAGAAAAGTTGCAAGACGACCAAGAGATTGAAACAAGAAAGCTGATACGAAAAAA 661
Db 1092 AAAAAAAAAAGAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1151
Qy 662 ATTGAGAAAGAAAAAGCAATGAGATATATATAGCAGAGATTTATATGTCGTTAG 721
Db 1152 AAGAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAG 1211
Qy 722 AAATACCAAGCTATAGAACTTCCATCAGAAAAATGAACTGATATTATATACCATCAAT 781
Db 1212 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAA 1271
Qy 782 CTCTCTTACCTCAGACACACAGAGGAAATAGTAGATTCACAGAAATCTATATAG 841
Db 1272 AAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1331
Qy 842 AAAAAACAATAGAGATCTATTACACAAATGTTGAAGACGAAAGGATATACATTAAG 901
Db 1332 AAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1391
Qy 902 GACATCTTGAAGAAAGAAAGATGGTTCAATTAACCAAGAAAGAAAGATTAATCTG 961
Db 1392 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAG 1451
Qy 962 CTGACATCAAAATCATCATTTAGAGACAGTAAATTTCTGATGTTAATGATTTTCAA 1021
Db 1452 AAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1511
Qy 1022 TAAGTAAATGATGAGATGAATAATGCTGAATATGACGATTAATGATGAAGAG 1081
Db 1512 AAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1571
Qy 1082 AAGATGATGAAGCTTGAAGCAATTTAAGCTATTTGCAATATGACAAATTTCCAGATG 1141
Db 1572 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1631
Qy 1142 AAGAAAAATGATGATTTATTAAGACACTAGAGATTTGATAGAGAAAAATGAATTTAG 1201
Db 1632 AGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGACAGAAAG 1691
Qy 1202 ATGATTTAGATGAAGAAATGAATAATCATCAGAGAAATTTATCGAAGAAAAATTA 1261
Db 1692 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1751
Qy 1262 AAGAAAAAATATGAAAAAACAAGATATATATTTTAAACCAATGATTAAGATTTGT 1321
Db 1752 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1811
Qy 1322 ATGATGACATATTAATAATAATATTAAGACAGTTAATAGGAAAAAGAAAAAT 1381
Db 1812 AAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1871
Qy 1382 TCATTAATCACTTCTTCAATATATTGACGAGACATGAATTTTCAAGTCGTGATG 1441
Db 1872 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1931
Qy 1442 AGTTATCTGAAGATATATATATATTTTATGAATACTATTA 1482
Db 1932 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1972

```

```

RESULT 7
DR141720 1932 bp mRNA linear EST 16-JUN-2005
LOCUS DR141720
DEFINITION 49300102 Drosophila pseudoobscura embryonic cDNA library Drosophila
ACCESSION DR141720
VERSION DR141720.1 GI:67886810
KEYWORDS EST.
SOURCE Drosophila pseudoobscura
ORGANISM Drosophila pseudoobscura
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

```

REFERENCE
AUTHORS
Neoptera: Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea: Drosophilidae; Drosophila.
1 (bases 1 to 1932)
Richards, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S.,
Nielsen, R., Thornton, K., Hubisz, M. J., Chen, R., Weisel, R. P.,
Couronne, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Bussemaker, H. J.,
van Batenburg, M. F., Howells, S. L., Scherer, S. E., Sodergren, E.,
Matthews, B. B., Crosby, M. A., Schroeder, A. J., Ortiz-Barrileños, D.,
Rives, C. M., Metzker, M. L., Muzny, D. M., Scott, G., Steffen, D.,
Wheeler, D. A., Worley, K. C., Havlak, P., Dudin, K. J., Egan, A.,
Gill, R., Hume, J., Morgan, M. B., Miner, G., Hamilton, C., Huang, Y.,
Waldron, L., Verdusco, D., Clerc-Blankenburg, K. P., Dubchak, I.,
Noor, M. A., Anderson, W., White, K. P., Clark, A. G., Schaeffer, S. W.,
Gelbart, W., Weinstock, G. M., and Gibbs, R. A.
Comparative genome sequencing of Drosophila pseudoobscura:
Chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)
15632085
JOURNAL
PUBMED
Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephen@bcm.tmc.edu
NCBI Trace Archive: 226717045
Insert Length: 1750 Std Error: 0.25.
FEATURES
source
1..1932
Location/Qualifiers
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="113"
/dev_stage="0-18h embryos"
/clone_lib="Drosophila pseudoobscura embryonic cDNA
library"
/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo
de priming from poly A+ RNA, directionally cloned"
ORIGIN
Query Match 16.4%; Score 243; DB 8; Length 1932;
Best Local Similarity 48.0%; Pred. No. 3.5e-30;
Matches 666; Conservative 0; Mismatches 722; Indels 0; Gaps 0;
Qy 2 AAGAACAAACGATCTGAAACAAGAGAGCTGCTAAGAAAGTTGCAAGAACAC 61
Db 251 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 310
Qy 62 AAGCGATTTAGAACAGATAGACTTGAAGAAAGTTTACAAAGACGCAAGCGATT 121
Db 311 AAAAAAAAAAAAAAAAAACACAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 370
Qy 122 TAGAACAGAGAGACTTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAAGAAC 181
Db 371 AAAAAAAAAAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 430
Qy 182 AGAGACGTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAGAGAGCTG 241
Db 431 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 490
Qy 242 CTAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAGATAGACTTGAAGAAA 301
Db 491 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 550
Qy 302 AGTTACAAAGACGCAAGCGATTTAAGAACAGAGAGCTGCTAAGAAAGTTGCAAG 361
Db 551 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 610
Qy 362 AACACAAAGCGATTTAAGAACAGAGAGAGCTGCTAAGAAAGTTGCAAGAACAA 421
Db 611 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 670
Qy 422 GCGATTTAAGAACAGAGAGACTTGCTAAGAAAGTTGCAAGAACAAAGCGATTTAG 481

```

```
Db      671 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 730
Qy      482 AACACAGAGAGCGTGTAAAGAAAAGTCCAGAAACAAGACGATTTAGACAAGAGA 541
Db      731 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 790
Qy      542 GACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGACAGAGACGTGCTA 601
Db      791 AAAAAAGAGAGAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAGA 850
Qy      602 AAGAAAAGTTGCAAGCAGCAAGAGATTGACAAACAGAAAGCGTATACGAAAAA 661
Db      851 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 910
Qy      662 ATTTGAAAGAAAAAGAAACATGAGATATATTGACAGAGATTATATGTCGTTAG 721
Db      911 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 970
Qy      722 AATATCCAGCTATAGAACTTCATCAGAAAAATGAAAGTGATTTATATACCATCAAT 781
Db      971 AAAAAAGAGAGAGAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAGA 1030
Qy      782 CTTCCTTACCTAGAGCAACAGAGGATATGAGATTCCAGAGAAATATCTATATAG 841
Db      1031 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1090
Qy      842 AAAAAACAAATAGAGATCTATTTACACAAATGTTGAGAGAGAGGATATACATAAG 901
Db      1091 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1150
Qy      902 GACATCTTGAGAGAAAAAGAGATGCTTCAATTAACAGAACAAAAAGAGATAATCTG 961
Db      1151 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1210
Qy      962 CTGACATCAAAATCATACATTAGAGACATTAATTTCTGATGTAATTTCAAA 1021
Db      1211 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1270
Qy      1022 TAGATAGATAGAGATGAATAATAGCTGAATATGACGATTCATTAATATGAGAG 1081
Db      1271 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1330
Qy      1082 AAGATGATAGAGCTTAGAGCAATTTAAGCTATTGTGCAATATGCAATTTCCAAGATG 1141
Db      1331 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1390
Qy      1142 AAGAAAACATAGGAATTTAAGAACTAGAAATTGATAGAGAAAAATGAAATTTAG 1201
Db      1391 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1450
Qy      1202 ATGATTTAGATAGAGATAGAAAAATCATCAGAGAAATTTCTGAGAGAAAAATTAAGA 1261
Db      1451 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1510
Qy      1262 AAGCAAGAAATATGAAAAAACAAGATATATTTTAACCAATGATTAAGATTGT 1321
Db      1511 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1570
Qy      1322 ATGATGAGCATATTAATAAATATGAATGATAGAGGTTATATAGAAAAAGAAAAAT 1381
Db      1571 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1630
Qy      1382 TCGTAAAA 1389
Db      1631 AAAAAA 1638
```

```
RESULT 8
CL114085 1727 bp DNA linear GSS 05-JAN-2004
LOCUS ISB1-59J11_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-59J11,
DEFINITION genomic survey sequence.
ACCESSION CL114085
```

```
VERSION CL114085.1 GI:40607720
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
REFERENCE Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Xenopodinae; Xenopus; Silurana.
AUTHORS Kremlitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the Xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 7500 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 280
High quality sequence stop: 1423.
FEATURES
source
location/Qualifiers
1..1727
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-59J11"
/clone_lib="ISB1"
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match 16.3% Score 241; DB 10; Length 1727;
Best Local Similarity 47.9%; Pred. No. 7.5e-30;
Matches 708; Conservative 0; Mismatches 769; Indels 1; Gaps 1;
Qy 5 AACACAAAGCATCTAGAACAGAGACGCTTAAGAAAAAGTTGCAGAACACAA 64
Db 128 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 187
Qy 65 GCGATTGACACAGATGACTTGTAAAGAAAGTTCAAGAGACAAAGCGATTAG 124
Db 188 TCTTCTTCAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 247
Qy 125 AACAGAGAGACCTTGAAGAAAGTGCAGAACACAAAGCGATCTAGAACAGAGA 184
Db 248 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 307
Qy 185 GACGTGCTAAAGAAAAAGTTGCAGAACACAAAGCGATTTAGAACAGAGACGTGCTA 244
Db 308 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 367
Qy 245 AAGAAAAGTTGCAGACACAAAGCGATTTAGAACAGATGACTTGTAAAGAAAGT 304
Db 368 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 427
Qy 305 TACAGAGACACAAAGCGATTTAGAACAGAGACGCTTAAGAAAAAGTTGCAGAGAC 364
Db 428 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 487
Qy 365 AACCAAGCGATTTAGAACAGAGACGCTTAAGAAAAAGTTGCAGAACACAAAGCG 424
Db 488 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 547
Qy 425 ATTTAGAACAGAGACCTTGAAGAAAGTTGCAGAACACAAAGCGATTTAGAAC 484
Db 548 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 607
Qy 485 AAGAGAGACGTGCTTAAGAAAAAGTTGCAGAACACAAAGCGATTTAGAACAGAGAC 544
Db 608 AAGAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 667
Qy 545 GTGCTTAAGAAAAAGTTGCAGAACACAAAGCGATTTAGAACAGAGAGACGTGCTAAG 604
```



```

QY 602 AAGAAAGTTGCAAGAGCAGCAAGAGATTGAAACAAGAGCGTGATACGAAAAA 661
Db 891 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 832
QY 662 ATTTAGAAAGAAAAAGAACATGAGATATATTAGCAGAGATTATATGCTGTTAG 721
Db 831 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 772
QY 722 AATATCCAGCTATGAACTTCATCGAAAAATGAGCTGGATTTATATCCACATCAT 781
Db 771 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 712
QY 782 CTTCCTTACCTCAGGACACAGAGGAGATAGAGATTCCAAGAAATATCTATATAG 841
Db 711 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 652
QY 842 AAAAAACAATAGAGAACTTATTACAACAATGTTGAGAGAGAGGATATACATAAG 901
Db 651 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 592
QY 902 GACATCTTGAAGAAAGAGATGTTCAATTAACCAACAACAAAAAGATTAATCTG 961
Db 591 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 532
QY 962 CTGACATCAAAATCATCATTAAGAGACAGTAAATTTCTGATGTTAATGATTTCAA 1021
Db 531 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 472
QY 1022 TAAGTAAGTATGAGATGAATTAAGTCTGAATATGACGATTCATTAAATGAGAAG 1081
Db 471 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 412
QY 1082 AAGATGATGAACATTAAGCAATTTAAGCTATTTGCAATATGACATTTCCAGATG 1141
Db 411 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 352
QY 1142 AAGAAACATAGGAATTTTAAAGAACTGAAGATTGATGAGAAAAATGAATTTAG 1201
Db 351 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 292
QY 1202 ATGATTTAGTGAAGATAGAAAAATCATCAGAGAAATTATCTGAGAAAAATTAATA 1261
Db 291 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 232
QY 1262 AAGGAAGAAATRTGAAAAACAAGATTAATTTTAACCAATGTTAAAGTTTGT 1321
Db 231 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 172
QY 1322 ATGATGAGCATATTAAAAATATTAATGATTAAGCAGGTTAATAGAAAAAGAAAT 1381
Db 171 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 112
QY 1382 TCATTAATCA 1392
Db 111 AAAAAAAAAANNA 101

```

```

RESULT 10
LOCUS DRI41924 1783 bp mRNA linear EST 16-JUN-2005
DEFINITION Drosophila pseudoobscura embryonic CDNA library Drosophila
DRI41924
ACCESSION DRI41924.1 GI:67887014
VERSION DRI41924
KEYWORDS EST.
SOURCE Drosophila pseudoobscura
ORGANISM Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1783)
AUTHORS Richardt,S., Liu,Y., Bettencourt,B.R., Hradecky,P., Letovsky,S.,

```

```

TITLE
JOURNAL
PUBMED
COMMENT
Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephenr@bcm.tmc.edu
NCBI Trace Archive: 226707030
Insert Length: 1750 Std Error: 0.25.
FEATURES
source
1..1783
Location/Qualifiers
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="07"
/dev_stage="0-18h embryos"
/clone_id="Drosophila pseudoobscura embryonic CDNA
library"
/notes="Vector: pOTB7, Site 1: EcoRI; Site 2: XhoI; o119o
dt priming from poly A+ RNA, directionally cloned"
ORIGIN
Query Match 16.1%; Score 238.4; DB 8; Length 1783;
Best Local Similarity 48.0%; Pred. No. 2e-29;
Matches 680; Conservative 0; Mismatches 738; Indels 0; Gaps 0;
QY 6 ACAACAAGCGATCTAGAACAGAGAGCGTCTTAAGAAAGTTGCAAGAACACAAG 65
Db 344 AAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 403
QY 66 CGATTGACAGCAAGTACCTTCTTAAGAAAGTTACAGAGCGCAAGCGATTTAGA 125
Db 404 AAAAAAAAAAGAGGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAA 463
QY 126 ACAAGAGACTTGCTAAAGAAAGTTGCAAGAACACAAGCGATCTAGAACAGAG 185
Db 464 AAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 523
QY 186 ACGTCTTAAGAAAGTTGCAAGAACACAAGCGATTTAGACAGAGAGCGTCTTA 245
Db 524 AAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 583
QY 246 AAGAAAGTTGCAAGAACACAAGCGATTTGAAACAAGATGATCTGTTAAGAAAGTT 305
Db 584 AAAAAAAAAAAGAAAAAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAA 643
QY 306 ACAAGAGCGCAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAGTTGCAAGACA 365
Db 644 AAAAAAAAAAAGGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 703
QY 366 ACAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAGTTGCAAGAACACAAGGGA 425
Db 704 AAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 763
QY 426 TTTGAACAAGAGACTTGCTTAAGAAAGTTGCAAGAACACAAGCGATTTAGACA 485
Db 764 AAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 823
QY 486 AAGAGAGCGTCTTAAGAAAGTTGCAAGAACACAAGCGATTTAGAACAGAGAGCG 545

```

```

Db      824 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 883
Qy      546 TGTAAAGAAAGTTGCAAGACAAAGCGATTGAAACAAGAGACGTCCTAAGA 605
Db      884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 943
Qy      606 AAAGTTGCAGACAGCAAGAAATTTGAAACAAGAAAGCGTGAATCGAATAAAATTT 665
Db      944 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1003
Qy      666 AGAAGAGAGAGAGAGAGATATATTAGCAGAGATTTATATGTCGTTAGAAAT 725
Db      1004 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1063
Qy      726 ACCAGCTATAGAACTTCCATCAGAAATGAACTGATATTATATACCATCACTTTC 785
Db      1064 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1123
Qy      786 TTTTACCTCAGGACACAGAGAGAAATAGTAAAGATTCACAGAAATATCTTATAGAAA 845
Db      1124 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1183
Qy      846 AACAAATAGAGAAATCTTATCAACAATTTGAAAGAGAGAGATTTACATTAAGACA 905
Db      1184 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1243
Qy      906 TCTTGAAGAGAGAGAGAGATGTTCAATTAACAGAAACAAGAGAGATTAATCTGCTGA 965
Db      1244 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1303
Qy      966 CATACAAATCTATCATTAAGACAGATTAATTTCTGATGTTATGATTTTCAATTAAG 1025
Db      1304 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1363
Qy      1026 TAAGTATGAGATGAATTAAGTGTGAATGACGATTCATTAATAGTAAAGAGAGA 1085
Db      1364 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1423
Qy      1086 TGATGAAGACTTAGACGAATTTAAGCCTATTTGCAATATGACAAATTTCCAAGATGAAGA 1145
Db      1424 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1483
Qy      1146 AAACATGAGATTTTATTAAGACATGAGAGATTTGATGAGAAAAATGAAAAATTTAGTGA 1205
Db      1484 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1543
Qy      1206 TTTAGATGAAGAGATTAAGAAATATGATCAGAGAAATTTATGAGAAAAATTAAGAAAG 1265
Db      1544 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1603
Qy      1266 AAGAAATATGAAGAGAGATTAATTTTAAACCAATGATTAAGATTTGTATGA 1325
Db      1604 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1663
Qy      1326 TGAGCATTTTAAATATTAATTAATGATTAAGCAGGTTAATTAAGAAAAAGAAAAATTCAT 1385
Db      1664 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1723
Qy      1386 AAATCATTTGTTTCATATATTTGACGAGCAATGAAGA 1423
Db      1724 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1761

```

```

RESULT 11
LOCUS   AJS92058      1434 bp      DNA      linear      GSS 15-JAN-2004
DEFINITION
Arabidopsis thaliana T-DNA flanking sequence, right border, clone
602D05, genomic survey sequence.
ACCESSION
AJS92058
VERSION
AJS92058.1 GI:37941682
KEYWORDS
GSS; right border; T-DNA flanking sequence.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana

```

```

REFERENCE
AUTHORS
1 Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chavain, S., Bechold, N., Cruaud, C., DeRose, R., Pelletier, G.,
Leplietec, L., Caboche, M. and Leclercq, A.
TITLE
T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
JOURNAL
EMBO Rep. 3 (12), 1152-1157 (2002)
PUBMED
12446565
REFERENCE
AUTHORS
2 (bases 1 to 1434)
Balzergue, S.
TITLE
Direct Submision
COMMENT
Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).
LOCATION/Qualifiers
source
1. 1434
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="602D05"
/clone_1ib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Wassilewskija"
1. 1434
/note="T-DNA flanking sequence
right border"
FEATURES
misc_feature
16.0%; Score 237.8; DB 10; Length 1434;
Best Local Similarity 44.5%; Pred. No. 2,5e-29;
Matches 617; Conservative 0; Mismatches 771; Indels 0; Gaps 0;
ORIGIN
Query Match 16.0%; Score 237.8; DB 10; Length 1434;
Best Local Similarity 44.5%; Pred. No. 2,5e-29;
Matches 617; Conservative 0; Mismatches 771; Indels 0; Gaps 0;
Qy      5 AACACAAAGCGATCTTAAGACAAAGAGACGTCCTAAGAAAGTTGCAAGAACAA 64
Db      1434 ANANNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1375
Qy      65 GCGATTTAGACAAAGATAGCTTGCTTAAGAAAGTTTACAGAGCAGCAAGGATTTAG 124
Db      1374 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1315
Qy      125 AACAGAGAGACTTGCTTAAGAAAGAGTTGCAAGACCAACAGCGATCTGAACAGAGA 184
Db      1314 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1255
Qy      185 GACGTCTAAGAGAAAGTTGCAAGAACCAACAAAGCGATTTGACAAAGAGAGACGTCCTA 244
Db      1254 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1195
Qy      245 AAGAAAAAGTTGCAAGACCAACAGCGATTTAGAACAGTTAGCTTGCTTAAGAAAGT 304
Db      1194 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1135
Qy      305 TACAGAGCAGCAAGCGATTTAGAACAAAGAGACGTCCTAAGAAAGTTGCAAGAAC 364
Db      1134 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1075
Qy      365 AACAAAGCGATTTAGAACAAAGAGACGTCCTAAGAAAGTTGCAAGAACCAACAGCG 424
Db      1074 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1015
Qy      425 ATTTAGAACAAAGAGACGTTGCTTAAGAAAGTTGCAAGAACCAACAGCGATTTGAGAC 484

```

```
Db      1014 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 955
Qy      485 AAGAGAGAGCTGCTTAAGAAAAGTTGCAAGAACAAAGCCATTATGACACAGAGAGC 544
Db      954 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 895
Qy      545 GTGCTTAAGAAAAGTTGCAAGAACAAAGCCATTATGACACAGAGAGCTGCTTAAG 604
Db      894 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 835
Qy      605 AAAAGTTGCAAGAGCAGCAAGAGATTGAACTAAAGAGAGCTGATCGAAAAAAATT 664
Db      834 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 775
Qy      665 TGAAGAAAGAAAAGAACATGAGATATTTAGCAGAGATTATATGTCCTTAGAAA 724
Db      774 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 715
Qy      725 TACCAAGCTATAGAACTTCATCGAATAATGAAAGTGTGATTTATATACCATCACTT 784
Db      714 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 655
Qy      785 CTTTAACTCAGAGCAACAGAGGGAATGATGAGATTCCAGGAATATCTATATAGAAA 844
Db      654 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 595
Qy      845 AACAACAATAGAAATCTATTACACAAATGTTGAAGAGAGAGGATATACATAAGAGC 904
Db      594 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 535
Qy      905 ATCTTGAAGAAAAGAGATGTTCAATATAAACCCAGAACAAAAGAGATAATCTGCTG 964
Db      534 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 475
Qy      965 ACATACAAATCATACATTAGAGACAGTAATATTTCTGATGTTAATGATTTCAATTA 1024
Db      474 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 415
Qy      1025 GTTATGATGAGATGAATAATGCTGATATGACATTCATTATATGATGAAGAGAG 1084
Db      414 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 355
Qy      1085 ATGATGAAGACTTGAAGCAATTTAAGCCTATTTGTCATATGACAATTTCCAAGATGA 1144
Db      354 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 295
Qy      1145 AAAACATAGAAATTTATAAGAACTTGAAGATTGATAGAGAAAATGAAAATTTAGATG 1204
Db      294 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 235
Qy      1205 ATTTAGATGAAGAAATGAAAATCTCGAAGAAATTAATCTGAGAAAATTAATAAAG 1264
Db      234 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 175
Qy      1265 GAAAGAAATATGAAAAAACAAGATATATTTTAAACCAATGATAAAGTTGTATG 1324
Db      174 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 115
Qy      1325 ATGAGCATTTTAAAAATATAAAAATGATAGCAGGTTATATAGAAAAGAAAAATTCA 1384
Db      114 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 55
Qy      1385 TAAATCA 1392
Db      54 AATTTAA 47
```

```
RESULT 12
CL078963      1808 bp      DNA      linear      GSS 31-DEC-2003
LOCUS      CL078963      1808 bp      DNA      linear      GSS 31-DEC-2003
DEFINITION      CH216-154E11, RM4.1 CH216 Xenopus tropicalis genomic clone
                  CH216-154E11, genomic survey sequence.
```

```
ACCESSION      CL078963
VERSION        CL078963.1  GI:40534876
KEYWORDS
SOURCE
ORGANISM      Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 1808)
Kremiczki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E., and Wilson, R.
A physical map of the xenopus tropicalis genome
JOURNAL
Unpublished (2003)
COMMENT
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 175000  Std Error: 0.00
Seq primer: RM4 ctcaaggcatcgctcgagc
Classes: BAC ends
High quality sequence start: 408
High quality sequence stop: 936.
Location/Qualifiers
1..1808
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-154E11"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTRABAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match      16.0%; Score 237.8; DB 10; Length 1808;
Beet Local Similarity 48.1%; Pred.No.2.5e-29;
Matches 668; Conservative 0; Mismatches 720; Indels 0; Gaps 0;

Qy      2 AAGAACAAACAAAGCATTTAGACACAGAGAGCTGTTAAGAAAAGTTGCAAGAACAC 61
Db      235 AAGAAANNAANNAANNAANNAANNAANNAANNAAGAAAGAAAAGAAAAGAAA 294
Qy      62 AAAGCATTTGAACAAGATGACTGCTAAGAAAAGTTTCAAGAGCAGAAACGATT 121
Db      295 GAAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 354
Qy      122 TAGAACAAAGAGACTTGCTTAAGAAAAGTTGCAAGAACAAACGATCTAGAACAG 181
Db      355 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAACAAAACAAAAAACA 414
Qy      182 AGAGAGCTGTTAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGAGCTG 241
Db      415 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAACAAAAAAGAAAAAAGAAAA 474
Qy      242 CTAAAGAAAAGTTGCAAGAACAAACGATTTAGAACAAAGTACTTGCTAAGAAA 301
Db      475 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 534
Qy      302 AGTTCAAGAGCAGCAAGCATTTAGAACAGAGAGAGCTGCTAAGAAAAGTTGCAAG 361
Db      535 AAAAAAAAAACAAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 594
Qy      362 AACACAAAGCATTTAGAACAGAGAGAGCTGCTAAGAAAAGTTGCAAGAACAAAC 421
Db      595 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 654
Qy      422 GCGATTTAGAACAGAGAGAGCTTGCTAAGAAAAGTTGCAAGAACAAACGATTTAG 481
Db      655 AAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 714
Qy      482 AACAAAGAGAGCTGCTAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAGAG 541
```

[illegible]

ORGANISM	Xenopus tropicalis
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.
AUTHORS	1 (bases 1 to 1654) Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE	A physical map of the xenopus tropicalis genome
JOURNAL	Unpublished (2003)
COMMENT	Contact: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@watson.wustl.edu Insert Length: 75000 Std Error: 0.00 Seq primer: T7 TAATACGACTCATCTATGAGG Class: BAC ends High quality sequence start: 468 High quality sequence stop: 812. Location/Qualifiers
FEATURES	Source 1..1654 /organism="Xenopus tropicalis" /mol_type="genomic DNA" /db_xref="taxon:8364" /clone="ISB1-72115" /clone_lib="ISB1" /note="Vector: pbeloBAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"
ORIGIN	Query Match 15.9%; Score 235; DB 10; Length 1654; Best Local Similarity 47.6%; Pred. No. 7,36-29; Matches 662; Conservative 0; Mismatches 726; Indels 3; Gaps 1;
OY	2 AAGAACAACAAGCGATTGTAGAACAAAGAGACGCTGAAGAAAAGTTGCAGAACAC 61
DB	29 AAAAACAACAAAAAACAABADAAGAAAAAGAAAGAAACATAAAAAAAAAAGA 88
OY	62 AAAGCGATTGAAACAAGATGACTTGCTTAAGAAAAGTTTCAAGAGCAGCAAGCGATT 121
DB	89 AAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAA 148
OY	122 TAGAACAAAGAGACTTGCTTAAGAAAAGTTGCAGAACAAAGCGATTGTAGACAG 181
DB	149 AAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAA 208
OY	182 AGAGACGCTGAAGAAAAGTTGCAGAACAAAGCGATTGTAGAACAAAGAGACGCTG 241
DB	209 AAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAA 268
OY	242 CTAAGAAAAGTTGCAGAACAAAGCG--ATTGAAACAAGATGACTTGCTTAAG 298
DB	269 AAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAA 328
OY	299 AAAAGTTTCAAGAGCAGAACGCGATTGTAGAACAAAGAGACGCTGAAGAAAAGTTGC 358
DB	329 AAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAA 388
OY	359 AAGAACAACAAGCGATTGTAGAACAAAGAGACGCTGAAGAAAAGTTGCAGAACAC 418
DB	389 AAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAA 448
OY	419 AAAGCGATTGAAACAAGAGACTTGCTTAAGAAAAGTTTCAAGAACAAACAAAGCGATT 478
DB	449 AAGAACAACAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 508
OY	479 TGAACAACAAGAGACGCTGAAGAAAAGTTGCAGAACAAAGCGATTGTAGAACAG 538
DB	509 AAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAA 568
OY	539 AGAGACGCTGAAGAAAAGTTGCAGAACAAAGCGATTGTAGAACAAAGAGACGCTG 598
DB	569 AAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAAAACAACAAAA 628

```

QY 599 CTAAGAAAAGTTGCAGAGCAGCAAGAGATTTAGAACAAAGAGCTGATACGAAA 658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 629 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 688
QY 659 AAAATTTAGAAAGAAAAGAAACATGAGATATATTAGCAGAGATTTATATGTCGT 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 689 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 748
QY 719 TAGAAATACCACTATAGAACTTCCTCAGAAAATGACGTGATTTATATCCACATC 778
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 749 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 808
QY 779 AATCTTCTTACCTCAGACAAACAGAGGATATAGATTTCAAGAAATATCTATTA 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 809 AAAAAAAAAAAAAAAAACCCCCAGACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 868
QY 839 TAGAAAAAAACAAATAGAGATCTATTACAACAAATGTTAGAGACGAGGATATACATA 898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 869 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 928
QY 899 AAGGACATCTTGAGAGAAAAGATGTTCAATTAACAGAACAAAAGATTAAT 958
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 929 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 988
QY 959 CTGCTGACATACAAATCATACATTAGAGACAGTAATATTTCTGATGTTAATGATTTTC 1018
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 989 AAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1048
QY 1019 AAATAGTAGTAGAGATGAATTAAGTCTGAATATGACGATTCATTATAGATGAAG 1078
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1049 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1108
QY 1079 AAGAGATATGAGACTTGAGAGATTTAAGCTATTGTCATATGCAATTTCCAG 1138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1109 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1168
QY 1139 ATGAGAAACATAGAGATTTTAAGAACTAGAAATTTGATAGAGAAAATGAATTT 1198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1169 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1228
QY 1199 TAGATGATTAGTAGAGATTAAGAAATCATCAGAGAAATTTATCTGAAGAAAATTA 1258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1229 AAAAAAATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1288
QY 1259 AAAAAAGAAAGAAATGAAAAAACAAGATTAATTTTAACCAATGATTAAGTT 1318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1289 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1348
QY 1319 TGTATGATGACATATTAATAAATATATTAATGATGAGGTTAATAGAAAAGATA 1378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1349 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1408
QY 1379 AATTCATAAA 1389
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1409 AAAACAGAAA 1419
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
LOCUS CL066150 1797 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-108B7 Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CL066150
VERSION CL066150.1 GI:40522063
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
          Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1797)
          Kremnitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
          Mardis, E. and Wilson, R.

```

```

TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
          Genome Sequencing Center
          Washington University School of Medicine
          Email: submissions@wustl.edu
          Insert Length: 175000 Std Error: 0.00
          Seq primer: Sp6 ATTAGGTGACACTATAG
          Class: BAC ends
          High quality sequence start: 306
          High quality sequence stop: 453.
          Location/Qualifiers
            source
              1..1797
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /strain="Nigerian frog"
                /db_xref="taxon:8364"
                /clone="CH216-108B7"
                /sex="male"
                /cell_line="Stock 248 F7A2, inbred N7"
                /clone_id="CH216"
                /note="Vector: pTRABAC2.1; CHORI-216 Xenopus tropicalis
                BAC library"

ORIGIN
Query Match 15.8%; Score 234.2; DB 10; Length 1797;
Beet Local Similarity 48.9%; Pred. No. 9.8e-29;
Matches 683; Conservative 0; Mismatches 705; Indels 10; Gaps 2;

QY 2 AAGAACAAACGATCTAGAACAGAGACGCTGAAGAAAAGTTGACAGAACAC 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 AAGAAAAAAACCCTAATAAAAAAGNNNANNNNAAAAAAAAAGAGAAAAA 201
QY 62 AAAGCATTTAGAACAAATAGACTTGTAAGAAATTTCAAGAGACGAAAGCAT 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 261
QY 122 TAGAACAGAGAGACTGCTAAGAAAAGTTGCAAGAACAAACGATCTAGAACAG 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 321
QY 182 AGAGACGCTGAAGAAAGTTGCAAGAACAAACGATTTGACAGAGACGCTG 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 AAGAAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 381
QY 242 CTAAGAAAAGTTGCAAGAACAAACGATTTGACAAAGTGAATGACTTGTAAGAA 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 441
QY 362 AGTTACAGAGCAGCAAGCGATTTAGAACAAAGAGACGCTGAAGAAAAGTTGCAAG 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 501
QY 362 AACCAACAAAGCATTTAGAACAGAGACGCTGAAGAAAAGTTGCAAGAACAA 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 561
QY 422 GCGATTTGACAGAGAGACTTGCTAAGAAAAGTTGCAAGAACAAAGCGATTTAG 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 AAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 621
QY 482 AACAGAGAGACGCTGAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAG 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 681
QY 542 GACGTCTAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAGACGCTGA 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 741
QY 602 AAGAAAAGTTGCAAGAGCAAGAGATTTAG-----ACAAAGAAAGCTGATACGA 656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 742 AAAAAAAAAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 801

```

Qy	657	AAAAAAAAATTAGAGAAAAAGAAACATGGAGATTATTAGACAGATTATATATGTCG	716
Db	802	AAAGAAAAAAGAAAAA	861
Qy	717	TTTAGAAATACAGCTATAGAACTTCCATCAGAAAAATGACGTGATTTATATACCA	776
Db	862	AAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	921
Qy	777	TCAATCTTCTTAACTCTCAGACACAGAGCGGATATGTAGATTCCAGAAATATCTAT	836
Db	922	AAAGAAAAA	981
Qy	837	AATGAAAAAAACAAATAGAAATCTATTATCAACAAATGTGAGACGAAAGGATATCA	896
Db	982	AAAGAGAAAAAA	1041
Qy	897	TAAAGACATCTTGAGAAAAAGAAAGATGCTTCAATAAACACAGAACAAAAAGATTA	956
Db	1042	GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	1100
Qy	957	ATCTGCTGACATACAAA-----TCATACATTAGACAGTAATATTTCTGATGTTAAT	1011
Db	1102	AAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAG	1161
Qy	1012	GATTTTCAAATTAAGTAACTATAGAGATGAATTAAGTCTGAATATGACGATTCATTATA	1071
Db	1162	AAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAG	1221
Qy	1072	GATGAGAAAGAAAGATGAGAGACTTAGACGAATTTAAGCCTATTGTGCAATATGACAT	1131
Db	1222	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAG	1281
Qy	1132	TTCCAGATGAGAAACATAGGAATTTATTAAGACTAGAGATTTGATAGAAAAAT	1191
Db	1282	AGAAAGAAACGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAG	1341
Qy	1192	GAAAAATTTAGATGATTGATGAGAGAAATGAAAAATCATCAGAAGAATTTCTGAGAA	1251
Db	1342	AAAAAAMCAAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	1401
Qy	1252	AAAAATAAAAAAGAAAGAAATATGAAAAAACAAAGGATTAATTTTAAACCAATGAT	1311
Db	1402	AACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAGAAAA	1461
Qy	1312	AAAGTTTGTATGATGACATTTTAAAAATTTAAATATGATTAACAGGTTTATTAAGAA	1371
Db	1462	AAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1521
Qy	1372	AAGGAAAAATTCATAAAA 1389	
Db	1522	AAAAAGAAAAAAGAAAAAAG 1539	
RESULT 15			
LOCUS	CL109409	1829 bp	DNA linear
DEFINITION	ISB1-51H21.Spe.1 ISB1 Xenopus tropicalis genomic clone ISB1-51H21,		
ACCESSION	CL109409		
VERSION	CL109409.1	GI:40603044	
KEYWORDS	SSS.		
SOURCE	Xenopus tropicalis (western clawed frog)		
ORGANISM	Xenopus tropicalis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;		
	Xenopodinae; Xenopus; Silurana.		
	1 (bases 1 to 1829)		
	Kiemlicki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,		
	Mardis, E. and Wilson, R.		
	A physical map of the xenopus tropicalis genome		
	Unpublished (2003)		
TITLE	Genome Sequencing Center		
JOURNAL	Contact: Richard K Wilson		
COMMENT			

FEATURES									
source									
<p>Washington University School of Medicine Email: submissions@wustl.wustl.edu Insert Length: 75000 Std Error: 0.00 Seq primer: Sp6 ATTATAGTGACACTAATG Class: BAC ends High quality sequence start: 856 High quality sequence stop: 936. Location/Qualifiers</p>									
1..1829									
/organism="Xenopus tropicalis"									
/mol_type="genomic DNA"									
/db_xref="taxon:8364"									
/clone="ISB1-51H21"									
/clone_lib="ISB1"									
/note="Vector: pbelOBAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"									
ORIGIN									
Query Match 15.8%; Score 234; DB 10; Length 1829;									
Best Local Similarity 47.7%; Pred. No. 1.1e-28;									
Matches 678; Conservative 0; Mismatches 742; Indels 0; Gaps 0;									
Qy	3	AGAACAA	CAAAAGCGATCTAGAAC	CAAGAGAGACGTCTAA	AGAAAAAGTTGCA	GAACAACA	62		
Db	153	AAAAGAA	AGAGNAAAAAAAAA	AGAAAAAAGAAAAA	AGAAAAAAGAAAAA	AGAAAAAAGAAAAA	212		
Qy	63	AAGCATTT	GAACAAGATAGACTTGCTAA	AGAAAAAGTTAC	AGAGCAAGCAAGCGATTT	122			
Db	213	AACAACCA	AAAAAGAAAAAGCGNCAG	AAAAATATAAAAA	AAAAAGAA	272			
Qy	123	AGAACAA	AGAGACTTGCTAAAGAAAAAGTTGCA	AGAACCAAGCGACTAGAAC	AGA	182			
Db	273	AAAAAAAA	AAAAAGAAAAAGAGAAAAAGAAAAA	AAAAAGAAAAA	AAAAAGAAAAA	332			
Qy	183	GAGACGT	GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAA	CAAGAGAGACGTGC	242				
Db	333	AAAAAAAA	AAAAAGAAAAAGAAAAA	AAAAAGAAAAA	AAAAAGAAAAA	392			
Qy	243	TAAAGAAA	AGTTGCAAGAACCAAAAGCGATTTAGAA	CAAGATAGACTTGCTAA	AGAAAA	302			
Db	393	GA	AAAAAAGAAAAAGAAAAA	AAAAATATAAAAAGAGAAAAA	AAAAA	452			
Qy	303	GTTACAA	GAGACGAAAGCGATTTAGAACAAAGAGACGTGCTAA	AGAAAAAGTTGCA	AGA	362			
Db	453	AAAAAA	GAACAAGAAAAAAGAAAAA	AAAAAGAAAAA	AAAAAGAAAAA	512			
Qy	363	ACAACA	AAAGCATTTAGAACAAAGAGACGTGCTAA	AGAAAAAGTTGCAAGAACAAAG	422				
Db	513	AAAAAA	GA	AAAAAAGAAAAA	AAAAAGAAAAA	572			
Qy	423	CGATTT	GAACAAGAGAGACTTGCTAAAGAAAAAGTTGCA	AGAACCAAGCGATTTAGAA	482				
Db	573	AAAAAA	AAAAAAGAAAAA	AAAAATATAAAAAGAAAAA	AAAAA	632			
Qy	483	ACAAGAG	AGACGTCTAAAGAAAAAGTTGCAAGAACCAAGCGATTTAGAAC	AGAG	542				
Db	633	AAAGAA	GAAGAAAAAAGAAAAA	AAAAAGAAAAA	AAAAAGAAAAA	692			
Qy	543	ACGTGCT	TAAGAAAAAGTTGCAAGAACCAAGCGATTTAGAAC	AGAGACGTGCTAA	602				
Db	693	AAAA	GA	AAAAAAGAAAAA	AAAAAGAAAAA	752			
Qy	603	AGAAA	AGTTGCAAGACGCAAGAGAGATTTAGAACAAAGAAAGCGTGA	TACGAAAAA	662				
Db	753	AGAAA	AAAAAAGAGAAAGAAAAA	AAAAAGAGAGAAAAA	CAAAAAA	812			
Qy	663	TTTAGAA	AGAAAAAGAACATGAGATATATTAGACAGAGATTTATATG	GTGCTTTAGAA	722				
Db	813	AAAA	AAAAAAGAAAAA	AAAAAGAAAAA	AAAAA	872			
Qy	723	AATAC	CGAGTATAGACTTCATGAGAAATGAACTGAGATATATTATAC	ACATCATC	782				
Db	873	AAAA	AAAAAAGAAAAA	AAAAAGAAAAA	AAAAA	932			

```
OY 763 TTCTTACCTCAGACAACAGAGGAATGATGAGATTCACAGAAATATCTATATATGA 842
Db 933 AAAAAATAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 992
OY 843 AAAAACAATAGAAATCTATTACACAAATGTTGAGAGCGAGGATATATCATTAAG 902
Db 993 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1052
OY 903 ACATCTTGAGAAAAAGAAATGTTCAATTAACCAAGACAAAAAGAAATTAATCTGC 962
Db 1053 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1112
OY 963 TGACATACAAATCATACATTAAGAGCAGTAATATTTCTGATGTTAATGATTTTCAAT 1022
Db 1113 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1172
OY 1023 AAGTAAGTATGAGATGAATTAAGTCTGAATATATGACGATTCATTAAATGATGAAGAGA 1082
Db 1173 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1232
OY 1083 AATGATGAAGACTTAGACGAATTTAAGCTTATGTCATATGACAAATTCAGAGATGA 1142
Db 1233 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1292
OY 1143 AGAAAACATAGGAATTTATTAAGAACTAGAGATTTGATAGAGAAAAATGAAAAATTTAGA 1202
Db 1293 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1352
OY 1203 TGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGATTATCTGAAGAAAAATTAAGAAA 1262
Db 1353 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAA 1412
OY 1263 AGGAAAGAAATATGAAAAAACAAGATTAATTTTAACCAATGATTAAGATTGTA 1322
Db 1413 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAA 1472
OY 1323 TGATGACATATTAATAAATATTAATAATGTAAGCAGGTTAATTAAGAAAAAGAAAAAT 1382
Db 1473 CAAGAAAAAGAAAAAGAAATGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAG 1532
OY 1383 CATAAATCATGTTTCATATATTTTGACGAGAGCAATGAA 1422
Db 1533 AAAAAAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1572
```

Search completed: December 7, 2005, 03:35:56
Job time : 6473 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 03:36:00 ; Search time 591 Seconds

(without alignments)
937.483 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1 CAGACACACACACACGATCT.....AATATTATGAACTATTA 1482

Sequence:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA-New:

```
1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB_seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB_seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB_seq.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB_seq.*
5: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB_seq.*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB_seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB_seq.*
9: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB_seq.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB_seq.*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77.6	5.2	169495	US-11-121-086-61	Sequence 61, App1
2	74.4	5.0	3181	US-10-793-626-386	Sequence 386, Ap
3	73	4.9	3543	US-10-793-626-4149	Sequence 4149, Ap
4	71.6	4.8	171486	US-11-121-086-105	Sequence 105, App
5	69.6	4.7	2655	US-10-793-626-1659	Sequence 1659, Ap
6	69.6	4.7	3520	US-10-793-626-4160	Sequence 4160, Ap
7	68.6	4.6	173602	US-11-121-086-25	Sequence 25, App1
8	68.4	4.6	173602	US-11-121-086-105	Sequence 105, App
9	67.6	4.6	3602	US-10-793-626-3582	Sequence 3582, Ap
10	67.6	4.6	15071	US-10-793-626-2963	Sequence 2963, Ap
11	67.2	4.5	3840	US-10-793-626-3187	Sequence 3187, Ap
12	67.2	4.5	4965	US-10-485-517-43	Sequence 43, App1
13	63.2	4.3	1434	US-11-082-746-2	Sequence 2, App1
14	63.2	4.3	6189	US-10-909-125-825	Sequence 825, App
15	63.8	4.2	627	US-10-793-626-693	Sequence 693, App
16	63.8	4.2	627	US-10-793-626-1325	Sequence 1325, Ap
17	62.8	4.2	2999	US-10-793-626-4339	Sequence 4039, Ap
18	62.8	4.2	3691	US-10-793-626-3338	Sequence 3338, App
19	62.8	4.2	173601	US-11-121-086-25	Sequence 25, App1
20	60.6	4.1	191091	US-11-121-086-60	Sequence 60, App1
21	59.4	4.0	3305	US-10-793-626-4002	Sequence 4002, Ap
22	59.4	4.0	150468	US-11-113-908-56	Sequence 56, App1
23	59.4	4.0	193789	US-11-113-908-55	Sequence 55, App1

C	24	59.2	4.0	2796	6	US-10-793-626-4335	Sequence 4335, Ap
C	25	58.8	4.0	1267	7	US-10-485-517-36	Sequence 36, App1
C	26	58.8	4.0	11896	6	US-11-121-086-84	Sequence 84, App1
C	27	58	3.9	139054	7	US-11-121-086-96	Sequence 96, App1
C	28	57.6	3.9	98862	7	US-11-121-086-76	Sequence 76, App1
C	29	57.6	3.9	182190	7	US-11-121-086-102	Sequence 102, App
C	30	57.4	3.9	207908	7	US-11-112-908-21	Sequence 21, App1
C	31	57.4	3.9	212805	7	US-11-112-908-19	Sequence 19, App1
C	32	57	3.8	151169	7	US-11-121-086-38	Sequence 38, App1
C	33	56.8	3.8	1752	6	US-10-793-626-1357	Sequence 1357, Ap
C	34	56.8	3.8	3055	6	US-10-793-626-4075	Sequence 4075, Ap
C	35	56.8	3.8	3819	6	US-10-131-826A-405	Sequence 405, App1
C	36	56.6	3.8	3509	6	US-11-077-550-19	Sequence 19, App1
C	37	56.4	3.8	1339	6	US-10-750-185-39902	Sequence 39902, A
C	38	56.4	3.8	120697	7	US-11-121-086-48	Sequence 48, App1
C	39	55.8	3.8	3162	6	US-10-793-626-4047	Sequence 4047, Ap
C	40	55.6	3.8	747	6	US-10-793-626-463	Sequence 463, App
C	41	55.6	3.8	2036	6	US-10-996-217A-8	Sequence 8, App1
C	42	55.6	3.8	2163	6	US-10-793-626-2057	Sequence 2057, Ap
C	43	55.6	3.8	3898	6	US-10-793-626-3962	Sequence 3962, Ap
C	44	55.4	3.7	3068	6	US-10-793-626-3824	Sequence 3824, Ap
C	45	55.2	3.7	2476	6	US-10-131-826A-489	Sequence 489, App

ALIGNMENTS

```
RESULT 1
US-11-121-086-61
: Sequence 61, Application US/11121086
: Publication No. US20050266459A1
: GENERAL INFORMATION:
: APPLICANT: FOUJSE, TIM S.
: APPLICANT: NIELSEN, KIRSTEN V.
: TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
: FILE REFERENCE: 09138.6000-00000
: CURRENT APPLICATION NUMBER: US/11/121,086
: CURRENT FILING DATE: 2005-05-04
: PRIOR APPLICATION NUMBER: 60/567,570
: PRIOR FILING DATE: 2004-05-04
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 61
: LENGTH: 169495
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: modified base
: LOCATION: (70072)..(70171)
: OTHER INFORMATION: a, c, g, t, unknown or other
: FEATURE:
: NAME/KEY: modified base
: LOCATION: (139457)..(157244)
: OTHER INFORMATION: a, c, g, t, unknown or other
US-11-121-086-61
```

Query Match 5.2%; Score 77.6; DB 7; Length 169495;
Best local similarity 45.1%; Pred. No. 0.0044;
Matches 287; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

Oy	3	AGAACACACACGATCTTAAAGAGAGACGCTTAAAGAAAGTTGACAGACACACA	62
Db	67358	ACGAC	67417
Oy	63	AAGCATTTTGAACACAGATGACTTGTAAAGAAAGTTTCAAGAGACACAAAGCATTT	122
Db	67418	GCACATCAACACACACACATCAACACACATCAACACACACACACACACACACACAC	67477
Oy	123	AGAACACACAGAGACTTCTTAAAGAAAGTTGACAGACACACAAAGCATCTTAAAGAC	182
Db	67478	TCAACACACACAGAGTAAACACACACACACACACACACACACACACACACACACAC	67537
Oy	183	GAGACGCTTAAAGAAAGTTGACAGACACACAAAGCATTTTGAACACAGAGACGTC	242

```
Db      67538 GATATCATCAAGCATGACGACCAACCAACAAAGCAAGCAACACACCACT 67597
Qy      243 TAAAGAAAAGTTGCAAGAAACAAGGCAATTTAGACAGATGCTTGAAGAAA 302
Db      67598 ACCACAAACAGCAACCACTCTACACAGCAACCAACCAACGCGCAACA 67657
Qy      303 GTTACAGAGCAAGGCAATTTGAAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGA 362
Db      67658 ACCACCAACCAACCAACAGCAACCAACCAACAGCAACCAACCAACAGCA 67717
Qy      363 ACAACAAAGCGATTTAGAAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAG 422
Db      67718 ACTACCAACCAACAGCAACGATTCATCAGCAACAGCAACCAACCAACCAACG 67777
Qy      423 CGATTTAGACAGAGAGAGCTTCTAAAGAAAAGTTGCAAGAACAAAGGCAATTTAGA 482
Db      67778 ACAACCAACAGCAACAGCATTCATCAGCAACAGCAACCAACCAACGCGCAACA 67837
Qy      483 ACAAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGACAGAGAG 542
Db      67838 ATATATGCAACAGCATTTCTAACACAGCAACCAACCAACGCAACCAAC 67897
Qy      543 ACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGACAGAGAGAGCGTCTAA 602
Db      67898 ACCTCAAAAACAGCAACCAACGAGAGCAATTAACAGCAACGAAGTGAACAATTA 67957
Qy      603 AGAAAAGTTGCAAGAGCAACCAAGATTTAGACA 638
Db      67958 GCAGCAATTAACAAAGCAGCACTGACATTAACCA 67993
```

RESULT 2

```
US-10-793-626-3866/c
; Sequence 3866, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3866
; LENGTH: 3181
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3866
```

```
Query Match      5.0%; Score 74.4; DB 6; Length 3181;
Beet Local Similarity 48.7%; Pred. No. 0.012;
Matches 266; Conservative 0; Mismatches 271; Indels 9; Gaps 2;
```

```
Qy      873 TGTGAGAGCGAAGGATATACATTAAGAGACTCTTGAAGAAAAGAGATGTTCAAT 932
Db      877 TATTAAAGAACTCCAGATGTACAGAAAGTAATACTAATGCTAATCTTAAGCTTCAAA 818
Qy      933 AAAACAGAAACAAAAGAAATTAATCTGTCATACATCAAAATCATCTTTAGAGAGT 992
Db      817 AACAACTTGTGCAATTTCAATTAAGCTTCAGATGTCMAAGAGTACGTGAGATTAATA 758
Qy      993 AATATTTCTGATGTTAA---TGATTTCAAAATTAAGTAATGAGATGAATA----- 1044
Db      757 GCAGAAATTTGATTAATAGCGTGAATTTGAAGAAATTTGCAAGGAAATGATCCAGCTTA 698
Qy      1045 -AGTGTGAATATGACGATTTCAATTAATAGATGAAGAAAGATGATGAAGACTTTAGACGA 1103
```

```
Db      697 CATTGAGAAAAAGCGAAAAATTAGCTTAAACAAAATGTAAGAGACGACAAATGGA 638
Qy      1104 ATTTAAGCTTATGTCATAATGCAATTTCCAGATGATGAAGAAAATAGGAAATTTATTA 1163
Db      637 CAAAAAGCTTCAAAAAGATTTAGAGAACTGCTATTAAGAGAAACAAAACAGCGTAAGC 578
Qy      1164 AGAATGAGAAATTTGATGAGAAAAATGAAAATTTAGATGATTTAGTGAAGAAATGA 1223
Db      577 TATACAGAAACAAAGTATTAAGACATGAAAAAACTCAAAAATTCGAAAGCAAGTTGG 518
Qy      1224 AAAATCATCAGAAATTTATCTGAAGAAAAATTAATAAAGAAAGAAATTTGAAAAAAC 1283
Db      517 CTTAATCTCTGTGTAATAATGATCGAAATTTGAGAAAAAGCGCAAACTAGAAAAAGA 458
Qy      1284 AAGGATTAATTAATTTTAACCAAAATGATTAAGATTTGTATGATGACATTTAAATAATA 1343
Db      457 TATGCTTAAGATTTTAAAAATTTAGATTAAGAGTTCAAAAGATTTGAAAAACGTC 398
Qy      1344 TAAAAATGATTAAGCAGTTAATTAAGAAAAAGAAAAATTCATTAATCATTGTTTCATAT 1403
Db      397 TAAAGAAAGACAAAACAGCGTGAAGAGCAGAGAAAGCTAGAAAAAAGAAATTTAAAAA 338
Qy      1404 ATTTGA 1409
Db      337 ATATGA 332
```

RESULT 3

```
US-10-793-626-4149/c
; Sequence 4149, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4149
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4149
```

```
Query Match      4.9%; Score 73; DB 6; Length 3543;
Beet Local Similarity 43.1%; Pred. No. 0.018;
Matches 581; Conservative 0; Mismatches 750; Indels 18; Gaps 4;
```

```
Qy      11 AAGGATCTAGAACAGAGAGACGTCCTAAAGAAAAGTTGCAAGAACAAAGCGATT 70
Db      3441 AAAAGAAATGCTAAAAAATATATGATTAAGAGCAGTGAATGAAGAAATTAATAATTCAAA 3382
Qy      71 TAGAACAGATAGACTTCTTAAAGAAAAGTTAACAAGACAGCAAGCGATTTAGAACAG 130
Db      3381 GAATTAATGATGCTACATATCAATTAAGAAAGCGAAATCGTAAAAATGAAAGAGCTA 3322
Qy      131 AGAGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAGAGCGTG 190
Db      3321 AGATTGAAGCAAAAGATATATTTCA-----ACGCAATGTACTAGAGATCAAGTAAATGA 3267
Qy      191 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTGAAGAAA 250
Db      3266 AGCGAAAATCTATAGTAATTAATAATTAAGAAATTAACACAGCAACTGATGTAATTC 3207
Qy      251 AGTTCAAGAACCAAAAGCGATTTTGAACAAGATGACTGCTTAAAGAAAAGTTTACAG 310
Db      3206 TGAAGCTAGACAGCAGTACAGAAATTAAGCAA-ATGAACAGATTAATCATATTCAAAACA 3148
```



```
QY 1242 ATCTGAAGAAAAATTAAGAAAAAGAAATATGAAAAAACAAGATTAATTATTA 1301
DB 16592 ATATATAATATATATATATATATATATATATATATATATATATATATATAT 16533
QY 1302 ACCAATGATATAAAGTTGTATGATGACATATTAATAAATATATAAATGATTAACAGST 1361
DB 16532 TGTTTATATATATATATATATATATATATATATATATATATATATATATATAT 16473
QY 1362 TAATAGAGAAAAAGAAAAATCATATAAATCATTTGTCATATATTTGACGAGACAAATGA 1421
DB 16472 AATCTTATATATATATATATATATATATATATATATATATATATATATATAT 16414
QY 1422 AATTTTACGATCGTGAGATGATTAATCTGAAGATTAATACTAAATTTTATGAAACTTAA 1481
DB 16413 ATTTATTAATATCAAAATATATATATATATATATATATATATATATATATATAT 16354

RESULT 5
US-10-793-626-1659
; Sequence 1659, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1659
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1659

Query Match 4.7%; Score 69.6; DB 6; Length 2655;
Best Local Similarity 43.9%; Pred. No. 0.051;
Matches 364; Conservative 0; Mismatches 454; Indels 12; Gaps 1;

QY 480 AGAACAAAGAGAGAGCTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACACAGA 539
DB 1383 AGGAGAGAGAGAGAGCTTAAACAAAAGTAATTTGTCACAAAAGATGAAATTTTCAACAT 1442
QY 540 GAGACTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACACAGAGAGAGCTGC 599
DB 1443 GCGAAATCACTTGAAGATTATTCACGAAACAGACGAAATTTGAACGTCAGCTTTAAAGA 1502
QY 600 TAAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCTGATACGAAAAA 659
DB 1503 ACAAACAAAACAGAGCTGAACATTAAGTGAACAAATTTTATGCGAAGTCAGCAAGTACA 1562
QY 660 AAATTTAGAAAAGAAAAGAGAAATGAGATATATTTAGAGAGAGATTTATATGTCGTTT 719
DB 1563 CAATTTAAAGAAACAAGTACATCATCAGAAATTTAGAACTGGA-----TAGACT 1610
QY 720 AGAAATACCAAGCTATAGAACTTCCATCAGAAAAATGAACCTGATTTATATACACATCA 779
DB 1611 AAAAACACAAAGAGCACTTTTAAATAATGAAACATGAAAGATTGAATTTGAAAAAATA 1670
QY 780 ATCTTCTTACCTCAGAGACACAGAGGATATAGATGATTCACAGGAATATCTTAAAT 839
DB 1671 TGGATATCAAAGTGAATTAAGTAAAGAAACATTTAAAGAAAAACAAATCATTTATTTGA 1730
QY 840 AGAAAAACAAATAGAGAAATCTATTTACAAACAAATGTTTGAAGACGAGAGGATATACATA 899
DB 1731 GATACAAACAAATAGAGCAACTAGAAAGTGAATTTGAAAGATATATACAAATTTATCAAA 1790
```

```
QY 900 AGCAGATCTTGAAAGAAAAAGATGGTTCATTAATAAACGAGACAAAAAGAGATTAATC 959
DB 1791 AGAAGAAAAAGCTTCGACACATCAACACAAACAACTACATCAAAAACATCGATTT 1850
QY 960 TGTGACATACAAAATCATCATATAGAGACAGTAAATATTTTCTGATGTTAATGATTTTCA 1019
DB 1851 AGCTGTTGTTTAAAGAGCAATTAATCGCAAAAGCAAGTTTATGACCTTTAGTAAACA 1910
QY 1020 AATAGTATAGATAGAGATGAAATAGAGTGAATGATGAGATTCATTTATATAGATGAGA 1079
DB 1911 ACTTAGCGATTCAGAGACCTCAAAAATTTGAAGTAAATGAAAAATCAAAATGTTTATTC 1970
QY 1080 AGAAGATGATGAAAGACTTAGACGAATTTAAGCTTATGCTAATGCAATATGACAAATTTCCAGA 1139
DB 1971 AGATGAATATGATGGGTAAAGATGCTTTTGAAGAAAGTGAAGAGCAAAATTCGACAAACAGA 2030
QY 1140 TGAAGAAAAATATAGAAATTTTAAAGAACTTAAAGATTTGATAGAAAAATGAAATTTT 1199
DB 2031 AAATGTAAAGCAAAATTTTAAATCAACAACTTATGATGATTAAGCGAACGTAAGATCT 2090
QY 1200 AGATGATTTAGATAGAGAAATGAAAAATCATCAGAAAGATTTATCTGAAAGAAAAATAA 1259
DB 2091 TAAATGAGAAATTCGAAATTAATGAAAGTCAAGCTTCAAAAGTCTATCAAGATATCTTTC 2150
QY 1260 AAAAGAAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATG 1309
DB 2151 TATGAAAAATCATTTATCAAGATATTTAAAGCAAAACAATCAAAAGCTAGATG 2200
```

```
RESULT 6
US-10-793-626-4160
; Sequence 4160, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4160
; LENGTH: 3520
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4160
```

```
Query Match 4.7%; Score 69.6; DB 6; Length 3520;
Best Local Similarity 43.9%; Pred. No. 0.051;
Matches 364; Conservative 0; Mismatches 454; Indels 12; Gaps 1;

QY 480 AGAACAAAGAGAGAGCTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACACAGA 539
DB 2248 AGGAGAGAGAGAGAGCTTAAACAAAAGTAATTTGTCACAAAAGATGAAATTTTCAACAT 2307
QY 540 GAGAGCTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCTGC 599
DB 2308 GCGAAATCACTTGAAGATTATTCAGACGAAACAGCAAAATTTGAACGTCAGTTAAAGA 2367
QY 600 TAAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCTGATACGAAAAA 659
DB 2368 ACAAACAAAACAGAGCTGAACATTTAAGTGAACAAATATTTTATGTCAGAGCAGCACTACA 2427
QY 660 AAATTTAGAAAAGAAAAGAGAAATGAGATATATTTAGACAGAGATTTATATGTCGTTT 719
DB 2428 CAATTTAAAGAAACAAAGTACATCATCAGAAATTTAGAACTGGA-----TAGACT 2475
QY 720 AGAAATACCAAGCTATAGAACTTCCATCAGAAAAATGAAAGTGAATTTATATACACATCA 779
```

Db	2476	AAAAACACAGACGACATCTTTAAAAATGAAATGAGAGATTGGATTTTGAAAAAATGGA	2535
Qy	780	ATCTTCTTTACCTCAGGACAAACAGAGGAAATGAGAGATTCGAAGAAATATCTATAT	839
Db	2536	TGGATATCAAAAGTGATTAAGATTAAGAAACATTTAAAGAAAAACAAATCATTTAAATGGA	2595
Qy	840	AGAAAAAACAAATAGAGAAATCTATTACACAAATGTTGAAGACGAAAGGATATACATA	899
Db	2596	GATCAACACAAATTTGAAGCAACTAGAAAGTGATTTGAAAAGATATACACATATTCAAA	2655
Qy	900	AGGACATCTTGAGAAAAAAGATGTTGTTCAATTAACCAACAAAAAAGAGTTAATC	959
Db	2656	AGAAAGAAAMCCTTCGACACATCAACCAACCAACTATCATAAAAACATCTGATTT	2715
Qy	960	TGCTGACATACAAATTCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCA	1019
Db	2716	AGCTGTTGTTAAAGCGCAATTTAAATCCGAAAAGCAGTTATATGACCTTTAGATTAACA	2775
Qy	1020	AATTAAGTAAATGATGAGATGAAATTAAGTCTGAATATGACGATTCATTAATAGATGAGA	1079
Db	2776	ACTTAGCATTCAGAACCTCAAAAAATTTGAAGTAAATGAAAAATTCAAATTTGTTTAATTC	2835
Qy	1080	AGAAAGATGATGAAGACTTAGACGAATTTTAACTTAATGTCAAATATGACAAATTTCCAGA	1139
Db	2836	AGATGAATATGATGGGTAAAGATGCTTTGAAAAGTTGAAGAAGCAAAATTCGACAAACAGA	2895
Qy	1140	TGAAGAAAACATAGCAATTTTAAAGAACTAGAACATTTGATAGAAAAATGAAAAATTT	1199
Db	2896	AAATGTAAAGACAAATTTTAAATTCACAACTTAGTGAGATTTTAAACGACGACGTAAGAATCT	2955
Qy	1200	AGATGATTTAGATGAGAAATAGAAAAATCATCAGAAATATTAATCTGAGAGAAAAATTA	1259
Db	2956	TAAATGAGAAATCGAAATTAATGAAAGTCAGCTTCAAAAATGTCATCAAGATATATCTTC	3015
Qy	1260	AAATGGAAAGAAATATGAAAAAACAAAGATTAATTAATTTAAACCAATTT	1309
Db	3016	TATGAAATATCATTTATCAAGATATTTAAAGCAAAACATCAAAAGCTAGATG	3065

	RESULT 7	
	US-11-121-086-25/c	
	Sequence 25, Application US/11121086	
	Publication No. US20050266459A1	
	GENERAL INFORMATION:	
	APPLICANT: POULSEN, TIM S.	
	INVENTOR: NIELSEN, KRISTEN V.	
	TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES	
	FILE REFERENCE: 09138, 6000-00000	
	CURRENT APPLICATION NUMBER: US/11/121, 086	
	CURRENT FILING DATE: 2005-05-04	
	PRIOR APPLICATION NUMBER: 60/567, 570	
	PRIOR FILING DATE: 2004-05-04	
	NUMBER OF SEQ ID NOS: 107	
	SOFTWARE: PatentIn version 3.3	
	SEQ ID NO 25	
	LENGTH: 173602	
	TYPE: DNA	
	ORGANISM: Homo sapiens	
	US-11-121-086-25	
	Query Match	
	Best Local Similarity 4.6%; Score 68.6; DB 7; Length 173602;	
	Matches 365; Conservative 0; Mismatches 449; Indels 7; Gaps 2;	
QY	656 AAAAAAAAAATTAGAAAGAAAAAGAACATGAGATATATTAGCAGAGATTATTCGTGC 715	
DB	136831 AATAATATATTAACAATTTATATTAATAATATATTAATATATTAACATATTTATATAATATA 136772	
QY	716 GTTTAGAATACCAGCTATAGAACTTCCTCATGAGAAAATGAACGTGATATTATATCCAC 775	
DB	136771 TAATATATTAACAATTTATATTAATAATATATTAATATTAATATTAATATAATATTAATATTAAT 136712	

Oy	776	TTCAATCTTCTTACCTCAGACAACAAGGAAATAGTAGATTCCAAAGAAATATCTA	835
Dd	136711	ATAATTAATATAT-----TATATTAATATATCTATATATAATATAGCATATTAATATATATA	136657
Oy	836	TAAATGAaaaaaaCAAATAGAGATCTATTACAAcAAATgTTGAAGGCGAAGGATATxG	895
Dd	136656	ATATATTTATTAATAAAATATATPAACATATTATTAATAAAATATATTAATAATATAAATTT--	136599
Oy	896	ATTAAGGCATCTTGAGAaaaaaGAAGTGGTCATATAAAACCGAACCAAAAAAGAGATA	955
Dd	136588	ATTAATATATTTATATATATTAATATATTTATATATACATTAACATATTAATAATATATATAT	136539
Oy	956	AATCTGCTGACATACAAAATCATACATTAGAGACGTAAATATTTCTGATGTAAATGATT	1015
Dd	136538	ACTATATAAAATATATATATTAATATTAATATATAAATATATACATATTTTAAATATAGTATA	136479
Oy	1016	TTCAATAATAGTAGATAGAGATGAATAAGCGCGAATATAGACATTCATTAATGATG	1075
Dd	136478	TATTTATATAGATTTTATATPAAATTTATATATATTTATATATATATATATATATATATATA	136419
Oy	1076	AAGAAGAAGATGATGAGMACCTTAGACGAATTTAAcCTATtTGCAATPAGCAAtTTGCC	1135
Dd	136418	TTTTATATATATATATATATATTTATATATATTTATATATATATATATATATATATATATG	136359
Oy	1136	AAGATGAAGAAAACTAGGAATTTTATAAGAACTGAAAGATTGGATPAGAAAAATGAAA	1195
Dd	136358	TATATATTTATATTAATAATAATAATTTATATATATATATATATAATATATATAATTTATATA	136299
Oy	1196	ATTATAGATGATTTAGATGAGAGATPAGAAAAATCATCGAAGAAATATCTGAGAaaaaa	1255
Dd	136298	TATTTATATATTTATATATAPAAATATATATATATTTATATATAATTTATATATATATATATA	136239
Oy	1256	TAAAAAAGAAAAAATATGAAAAAACAAAGCATATATTTTAAACCaaaTGATAAAA	1315
Dd	136238	TATATAAATTTTATATATTTATATATATATATATATATTTATATATATATATATATATATTT	136179
Oy	1316	GTTTGATGATGAGCATTTTAaaaaaATTAaaaaAGATAAGCGTTTAAAGAAAGG	1375
Dd	136178	TATATTTATATATCTATATATAATATATATATTTATATATTTTATATATTAATATATATATNT	136119
Oy	1376	AAAAATTCATAAAATCATTTGTTTCATATATTTGACGAGACAGATGAATTTTACAGATCG	1435
Dd	136118	ATATTTATATATAAAAATATATACADTATATTAATATATATATATATTTACATATACATATPA	136059
Oy	1436	TGATGAGATTATCGAAGATATACTTAAATATTTATGTAAA	1476
Dd	136058	AATATATGTAAAAATGTATGTATGATGAATATATGTATATATATA	136018

```

RESULT 8
US-11-121-086-105
; Sequence 105, Application US/11121086
; Publication No. US2005026459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 105
; LENGTH: 171486
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-105

Query Match          4.6%; Score 68.4; DB 7; Length 171486;
Best Local Similarity 46.2%; Pred. No. 0.075;
Matches 370; Conservative 0; Mismatches 421; Indels 9; Gaps 4;

```

Oy		681	ACATGAGATTATTTAGCAGAGATTTTATGTGGCTTAGAAATTCAGCGTAAGAAT	740
Db		16086	AAAATATAAATATAATATAATATAATATAATATAATATAATATAATATAATATA-T	16144
Oy		741	TCCATTCAGAAAATGAACGTGGATATTATATATACACATCATCTTCTTACTCAGACAA	800
Db		16145	ATTATATATATATATATCTTATATATTTATATATATATATATATATATATATATATATAT	16204
Oy		801	CAGAGGAATATGATGAGATTCCAAGCAATATCTATATATGAAAAACAATATGAGATC	860
Db		16205	TAT	16264
Oy		861	TATATACAAACAAGCTTGAGAGAGGAGATATACATTAAGGCATCTTGAGAAAAAGA	920
Db		16265	TTAT	16318
Oy		921	AGATGTTTCATATAAACACAGAACAAAAAGAAAGATTAATCTGTCGATACAAATCATAC	980
Db		16319	AATATTTAT	16378
Oy		981	ATTAGACACGTAATAATATTTCTGATGTTAAT-GATTTCAAATATAGTAATGATGAGATG	1039
Db		16379	ATTAT	16438
Oy		1040	AAATATAGTCGATATATGAGATTCATTAATATAGATGAGAGAGATGATGAGAACTTGA	1099
Db		16439	TTTAT	16498
Oy		1100	ACGAATTTTAAAGCCATATGTCGCAATATGACAAATTTCCAGATGAGAGAAAACATAGAA	1159
Db		16499	ATTAT	16558
Oy		1160	ATA-AAGACTGAGAGATTGATATGAGAAAAAGAAATTTAGATGATTTAGATGAGAGA	1218
Db		16559	AT	16618
Oy		1219	ATAGAAAAATCATCAGAGAAATPATCTGAGAAAAATATAAAAAGAAAGAAATATGAA	1278
Db		16619	TTTAT	16678
Oy		1279	AAAACAAGATTAATATTTTAAACCAATGATMAAGTTTGATGATGAGCATATTTAA	1338
Db		16679	AAT	16738
Oy		1339	AAAT	1398
Db		16739	TAAAT	16798
Oy		1399	CATATATTTGCGGAGACAATGAAATTTTACAGATCGGATGAGTATCTGAGAAATATA	1458
Db		16799	TAAAT	16858
Oy		1459	ACTAATATATTTATGAACCT	1478
Db		16859	TATTTATATATATATATATATAT	16878
RESULT 9				
US-10-793-626-3582/c				
; Sequence 3582, Application US/10793626				
; Publication No. US20050255478A1				
; GENERAL INFORMATION:				
; APPLICANT: KIMBERLY, WILLIAM JOHN				
; TITLE OF INVENTION: STRAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS				
; FILE REFERENCE: PUS480US				
; CURRENT APPLICATION NUMBER: US/10/793.626				
; CURRENT FILING DATE: 2004-03-04				
; PRIOR APPLICATION NUMBER: 60/164,258				
; PRIOR FILING DATE: 1999-11-09				
; NUMBER OF SEQ ID NOS: 4472				
; SOFTWARE: PatentIn Ver. 2.1				
; SEQ ID NO 3582				

```

; LENGTH: 3602
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3582

```

Query Match	4.6%	Score 67.6;	DB 6;	Length 3602;
Best Local Similarity	43.1%	Pred. No. 0.095;		
Matches 438; Conservative	0;	Mismatches 569;	Indels 9;	Gaps 2;

QY	333	AGAGAGACGTGCTAAAGAAAGTTGGCAAGAACAAAGAGATTTAGAACAGAGACG	392
Db	2208	AGTAAATCCGTGTAAACCAAGTTAAATTTAGAAAGCGTGATTAATTTAAACGTCA	2149
QY	393	TGCTAAAGAAAGTTGGAGAACCAACCAACGATTTTGAAACAAGAGACTGTGTAAGA	452
Db	2148	ATTTAACCAAGCTTTAAAGAGAGCTAAGGGGATTAATTAATAATATGCGTACAAATGT	2089
QY	453	AAAGTGCAGAGAACAAACCAAGCGATTTTGAAACAAGAGACGTGCTAAAGAAAGTTGCA	512
Db	2088	CAATATCATATGACATTCACATCTTTTAACCAAAATGATATATGCTTAAGCCAAATTAAA	2029
QY	513	AGAAACAACAAAGCATTTTGAACAAGAGACGTGCTAAAGAAAGTTGCAACAACA	572
Db	2028	TGGTGAACGACGTCTAAAGAACATCAACAAAATCTGAAGTATTTATTTAAAGAAAT	1969
QY	573	AAGGATTTTGAACAAGAGACGTGCTAAAGAAAGTTGCAAGAGACCAAGAGATTT	632
Db	1968	AGATATACCTTAATATATGCTCAAAAAGCTGCATTA-----TTATATGATTTAGAGGCTC	1915
QY	633	AGAACAAAGAAAGCGTGAATCGAAGAAAAAAATTTAGAAAGAAAAAGGAAACATGAGATAT	692
Db	1914	TAAAGACATTAATAATTAATTAATCAAAATCGTTGATTAATGCAATAGAAATTAATATGATGCTAT	1855
QY	693	ATTAGCAGAGATTTATATGTGCTGTTTGAATAACGACATATAGAACTTCATCAGAAA	752
Db	1854	GCAAGGTTTAAAGAACATGTAGCTCAATTAACAGCAACTTACAAAMACAAACATTTGATATA	1795
QY	753	TGAACGTGGAATTAATATATACACATCAATCTCTTACCTCAGGACAAAGAGGAATAG	812
Db	1794	TTTAATGTGATGAGACCTTAATAATATCAATATGATTAGCTATCAACTTAGGCAATTA	1735
QY	813	TAGAGATTCGAAGAAATATCTATATATAGAAAAAAACAATATGAGAAATCTATTACAACAA	872
Db	1734	TGTTCTTGACAAAGAAAACGCTACAAATTAAGAAGCGTAATATCAATATGGAATGATTC	1675
QY	873	TGTTGAMGACGAAGGGATATATCATTAAGAATCTTGGAAGAAAAAGATGSGTTCAAT	932
Db	1674	AAACATGATGATGCTAGAGACATCTTAATATGAAATGGAAGACTTTAAAGTGTCCAAC	1615
QY	933	AAACCCGAAACAAAAGAGTAATATCGTGAATCAACAAATCATATATAGAGACAGT	992
Db	1614	AAAAGCATATATGACATTAATAGATAGCTCAACGTCACACTTATGAATTTGAACACGC	1555
QY	993	AAATATTTCTGATGTTATATGATTTTCAAAATAGATAGATAGAGATGAATATAGTCTGA	1052
Db	1554	TAAAGCAACATCAAAATCTTAAGCTCAAGCAAAAACAATGGTAAATAGGAAA--GCTAG	1498
QY	1053	ATATGACATTCATTAATATAGATGAAGAAAGATGATGAAGACTTAGACGAATTTAAAGC	1112
Db	1497	AAAAGCGTTTCTATATTTAATATCAAGCAACATCAAAATGATTTAGTTATCAAGCAAAA	1438
QY	1113	TATGTGCAATATGACATTTCCAAGATGAAGAAAACATAGGAATTTATAAAGAACTAGA	1172
Db	1437	TGAAGGGCAATCTGCAATTTGAACATATATGATGCAAGATTAATTAAGCAAAACCTTGA	1378
QY	1173	AGATTTGATAGAAAAATGAAAAATTTAGATTTTATAGATGAGAAATAGAAAAATCATC	1232
Db	1377	TGCTATCAATATGATGACAAAAGTTGAAGATATATAATCATCTTAATATGCAAAATCC	1318
QY	1233	AGAAAGATTAATCTGAAGAAAAATTAABAAAAGAAAGAAATATGABAAAAACAAAGATTA	1292

Db 1317 AAACCTATCAAAAGAGAAAAAATAAATACTAATCTCAAAATTAATAGTAAGTAATG 1258
Qy 1293 TAAATTTAAACCAATGATAAAGTTGTATGATGACGATATTAATAATATATAA 1348
Db 1257 AATTAAAGATGAATTCACAGAGTATTAACCAACCAATAGAAAATGCTACAA 1202

RESULT 10
US-10-793-626-2963
; Sequence 2963, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; FILE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; TITLE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2963
; LENGTH: 15071
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2963

Query Match 4.6%; Score 67.6; DB 6; Length 15071;
Best Local Similarity 43.1%; Pred. No. 0.095;
Matches 438; Conservative 0; Mismatches 569; Indels 9; Gaps 2;

Qy 333 AGAAGACGTCCTAAAGAAAGTGCAGAACCAACAAAGCATTTAGAACAGAGACG 392
Db 12864 AGTAATCTCGTTAAAGCAAGTTTAATTAAGAAAGCGTATTAATTAATAAGCTCA 12923
Qy 393 TGCTAAAGAAAGTTGCAGAACCAACAAAGCGATTTAGAACAGAGACCTGCTAAAGA 452
Db 12924 ATTTAACAAAGCTTTAAAGAAAGCTAAAGCGTATTAATTAATAATGCGGTACAAATGT 12983
Qy 453 AAAGTTGCAGAACCAACAAAGCGATTTAGAACAGAGACGCTTAAGAAAGTTGCA 512
Db 12984 CAATATCAATGACATTCATCAATCTTTTAACAAATAGATATGCTAAAGCCATTAATA 13043
Qy 513 AGAACCAACAAAGCGATTTAGAACAGAGACGCTTAAGAAAGTTGCAGAACCAACA 572
Db 13044 TGTGGAACGACGCTTAAGAAAGCAATCAACAAAAATCTGAAGTATTTATTAAGAATT 13103
Qy 573 AAGGATTTAGAACAGAGACGCTTAAGAAAGTTGCAGAACCAAGATT 632
Db 13104 AGATATCTTAATATGCTCAAAAAGCTGCAATTA-----TTAATCAGATTAAGGCTC 13157
Qy 633 AGAACCAACAAAGCGTATGACAAATAATTTAGAAAGAAAGAAAGCAATGAGATAT 692
Db 13158 TAAAGACATTAATAATTAATTAATCAAAATGCTGATATGACAAATGAAATTAATGCTAT 13217
Qy 693 ATTACAGAGATTTATATGCTGCTTTAGAAATACAGCTTAAGAACTTCATCAGAAA 752
Db 13218 GCAAGGTTAAAGCAACATGATGCTCAATTAACAGCACTCAAAAGCAACATTTGATA 13277
Qy 753 TGAACGTGATATTAATATACCAATCAATCTTCTTACCTGAGCAACAGAGGAAATAG 812
Db 13278 TTTAATCTGATGAGACTTTAAATATCAATATGATTAATGCTATCACTTAAGGAATTA 13337
Qy 813 TAGAGATTCAGAGCAATATCTATATAGAAAAAACAATAGAGATCTAATTAACAACA 872
Db 13338 TGTTCCTTACAAAGAAAAGGATGACAAATTAAGAGCTAATATCAATATGGAATGATCA 13397
Qy 873 TGTGAAGAGCAAGGATATATCATTAAGACATTTTGAAGAAAGAAAGTGTTCAT 932

Db 13398 AAACATGATGATGCTAGAGACCTCTTAATGAAATTGAAGACTTAAGATGCTCAAC 13457
Qy 933 AAACCGAGAACAAAAGAGATTAATCTGCTGACATACAAATTCATATGAGACAGT 992
Db 13458 AAAAGCATAATGACATTAAGATGCTCAAGCTCAACTTGATGAATTAAGACAGC 13517
Qy 993 AAATATTTCTGATGTTAATGATTTTCAATAAGTAAGTAAGATGAAATTAAGCTGA 1052
Db 13518 TAATGCAATCAAAATTCCTTAAGCTCAGCAAAAACAATGTAATGAGAA--GCTAG 13574
Qy 1053 AATATGCAATTCATTAATGATGAGAGAGATGATGAAGCTTAAGCAATTTAAGCC 1112
Db 13575 AAAAGGTTTCTTAATATTAATACGCAACATCAATGATTTAGTTAAATCAAGCAAAAGA 13634
Qy 1113 TATGTCATATGACAAATTTCCAGATGAAGAAACATAGGATTTATTAAGAACTAGA 1172
Db 13635 TGAAGGCAATCTGCAATTAAGACATACATGAGATGAAATTAATCTTAAGCAAACTAGA 13694
Qy 1173 AGATTTGATAGAGAAAAATGAAATTTAGATGTTTGAATGAAGAAATGAATCATC 1232
Db 13695 TGCTAATCAATGATGACCAAAAAGTTGAAGATTAATCACTTAATTAAGTCAAAATGC 13754
Qy 1233 AGAAGATTAATCTGAAGAAAAATTAATAAGAAAGAAATATGAAAAACAAGATTA 1292
Db 13755 AAACCTATCAATGAGAAAAAATAAATTAATCTCAAAATTAATTAAGTTAATGTAATG 13814
Qy 1293 TAAATTTAAACCAATGATTAAGTTGTTGATGATGACATTAATTAATAATTAATA 1348
Db 13815 AATTAAAGATGAATTCACACAGCTATTAACAAACAAATTAAGAAATGCTACAA 13870

RESULT 11

US-10-793-626-3187
; Sequence 3187, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; FILE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; TITLE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3187
; LENGTH: 3840
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3187

Query Match 4.5%; Score 67.2; DB 6; Length 3840;
Best Local Similarity 43.4%; Pred. No. 0.11;
Matches 425; Conservative 0; Mismatches 543; Indels 12; Gaps 2;

Qy 380 AACAGAGACGTCCTTAAGAAAGTGCAGAACCAACAAAGCGATTTAGAACAGAGAGA 439
Db 137 AAGCACAATTAATGCAAGACATACCAACCAAGGTGTCAAATCTTAACAAAGAGCGCA 136
Qy 440 GACTTGCTAAAGAAAGTTGCAGAACCAACAAAGCGATTTAGAACAGAGACGCTGA 499
Db 197 TAACCTCTTATATCTCGAATTAATGCAACAAAGTTGTTGGAAGAGCTGCAAGAAATGCA 256
Qy 500 AAGAAAGTTGCAGAACCAACAAAGCGATTTAGAACAGAGACGCTGAAGAAAGT 559
Db 257 TCGAAACAAAGGCAACCAACAAAGCGATTTATTAATTAATA--ATGATTAATGCTACAG 313
Qy 560 TGCAGAGAACCAAGGATTTAGAACAGAGACGCTGCTAAAGAAAGTTGCAAGAGC 619
Db 314 ATGAAGAAAAGAGGCTGCCAACATTTAGTTATCGCTACAAAACAAATATCATTTAGATA 373

Oy	620	GGCAAGAGATTTTGAACAAGAGAGCGTATACGAAAAAAATTTAGAAAAAGAAAG	679
Db	374	ATATTACTCCTTATCTTCAATATATGATGTTGAAAAAGCTTAAGTAGCAGGAATTAATG	433
Oy	680	AACATGAGATATATATTAGCAGAGATTTATATATGTCGTTTAGAAATACAGCTATAGAAC	739
Db	434	AAATAGCTAAAGTTTTTACAGAAACCGCTGTTTAAGTCAAAAGCAAAAAAGATATTTGATC	493
Oy	740	TTCCATCGAAAAATGACAGTGGATATTATATACCATCTAATCTTCTTACTCAGAGCA	799
Db	494	AAAACTGCGCCACAGATTAATCAAAATTCAAACGCACTCAACCTGCTCAACTGCGGAAA	553
Oy	800	ACAGAGGAAATAGTAGAGATTCGAAAGAAATATCTATATAGAAAAAAACAATAGAGAT	859
Db	554	AAGAAAGCGGCTATTCAATTTGGCAATCAAAAATCAATAGAAAGCAAGAACAGCAATTCAA	613
Oy	860	CTATTACAACAATGTTTGAAGGACGAAAGGATATACATPAAAGACATCTTGA-----	911
Db	614	ATGAACATATATTAACAATGGTGTGCGCAAGCTTAATCTAACGGCATTCATGAATTTGAAT	673
Oy	912	-AGAAAAAGAAAGATGTCCTCAATTAACCAAGACAAAAAGAAAGATTAATCTGCTGACATAC	970
Db	674	TAGTTATCCAGATGCGCAAAAAATCTGATGCTPAAACAAGTATCGATATATTAATATA	733
Oy	971	AAAAATCATACATTTAGAGACAGTAATATTTCTGATGTTAATGATTTTCAAAATTAAGTAAGT	1038
Db	734	ATGAGCAAAATATATCTATCAACACTACACGATGCAACAGATGAAGAAAAAGCAAAAAG	793
Oy	1031	ATGAGGATGAATAATAGTCTGTAATATGACATTCATTAATATGATGAAGAGATGATG	1090
Db	794	CATTAGATTAATTAATAAATAGCTTAAGATGACAGATACAAACAAGTTGATCAAGCGCAAA	853
Oy	1091	AAGACTTGAACGAATTTAAAGCTATATGCGCAATAGACATTTCCAAATGAAGAAACA	1155
Db	854	CAAAACCAACAAGTATCTGATGCAAAAACCTGAGGCTATAGATACGATTAATAATTCAAG	913
Oy	1151	TAGCAATTTTATTAAGAACTAGAAAGATTTGATGAGAAAAATGAATAATTTAGATTTAG	1210
Db	914	CAAAATGTTGCAAAAAAACCATCCGCTCGAGTGGATTTGATTCAAAAGTTTGAGATTTTAA	973
Oy	1211	ATGAAGAAATGAAAAAATCATCAGAAAGATTATCTGAAAGAAAAATAAAAAGAAAGA	1270
Db	974	AGCGTCATCAATCAATCAACGCCCATGCTACAGAAAGAAAAACAAGATGCAATTTCAA	1033
Oy	1271	AATATGAAAAAACAAGATTAATATTTTAAACAAATGATPAAAGTTTGTATGATGAC	1330
Db	1034	GATTGAAATGTTAAAGATGAAGTTTAAAGATCTAATTAATCAAGATGACGTCGAATG	1093
Oy	1331	ATATTAAAAAATATATAAAT 1350	
Db	1094	AAATTGAACAGCAAAAAAT 1113	
RESULT 12			
US-10-485-517-43/c			
; Sequence 43, Application US/10485517			
; Publication No. US20050256299A1			
; GENERAL INFORMATION:			
; APPLICANT: University of Sheffield			
; APPLICANT: Biosynexus Incorporated			
; APPLICANT: Foster, Simon			
; APPLICANT: Mond, James			
; TITLE OF INVENTION: Antigenic Polypeptides			
; FILE REFERENCE: P100629MO			
; CURRENT APPLICATION NUMBER: US/10/485,517			
; CURRENT FILING DATE: 2004-02-02			
; PRIOR APPLICATION NUMBER: GB 0118825.9			
; PRIOR FILING DATE: 2001-08-02			
; PRIOR APPLICATION NUMBER: GB 0200349.9			
; PRIOR FILING DATE: 2002-01-09			
; NUMBER OF SEQ ID NOS: 424			
; SOFTWARE: PatentIn version 3.1			

```

; SEQ ID NO 43
; LENGTH: 4965
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-43

Query Match      4.5%; Score 67.2; DB 6; Length 4965;
Beet Local Similarity 45.1%; Pred. No. 0.11;
Matches 341; Conservative 0; Mismatches 403; Indels 12; Gaps 2;

OY      563  AAGAACCAACAAACGATTTAGAACAGAGAGACGTCCTAAAGAAAAGTTGCAGAGCAGC 622
      |||||
DB      3843  AAGAGAAAAATGACAGAAAAAGATTAGTTTAAAGCTAAAGAAAAAGCTATCAAGATA 3784

OY      623  AAAGGATTTAGAACAAAGAAAGCGCTATCCGAAAAAAATTTAGAAAGAAAAAGAAC 682
      |||||
DB      3783  TCTTAAATGCAACAAACMACTAAATGATGTACGCAATTTAAAGATCAAGCAGTTCCTATA 3724

OY      683  ATGGAGATATATTAGACAGAGATTTATATATGTCGTTAGAAATACCGATATGAACTTC 742
      |||||
DB      3723  TTCAAGGTATTACTGCAATACAACTAATTAAGATTTGGAAAGATGAATTAGCAACAA 3664

OY      743  CATCAGAAAATGAGCGTGATATTAATATACCATCATCTCTTTTACCTCAGACACACA 802
      |||||
DB      3663  AAGCAAAACGACAAAAGCGCTTATTCGACAACTGCAGATGGCAGTACAGGAAAAAG 3604

OY      803  GAGGAAATGTAGAGATTCGAAGAAATCTTAAATAGAAAAACAATAGAGATCTA 862
      |||||
DB      3603  AACCAAGCAAAATCAACAGTAGACGCACAAATTAACAACAGGTAATCAAAATATGMAAATG 3544

OY      863  TTACAAACAAATGTTAGAGACGAAAGGATATCATTAAGAGACATCTTGAAGAAAAAGAA 922
      |||||
DB      3543  CACAGTCAATCGATGATGTGAAACACTGCAGAAAGATATGCAATTCAGCAATTGACCCAA 3484

OY      923  ATGTTTCATTAATAACCGAAGCAAAAAGAAATMACTCG-----CTGACATACAAA 973
      |||||
DB      3483  TTCAAGCATCAACAGATGTTTAAACGAATGCAGAACGGAAATGCTTAAGTGAATGCAAA 3424

OY      974  ATCATACATTTAGAGACAGTAAATATTTCTGATGTATATGATTTTCAATATAGTAGATG 1033
      |||||
DB      3423  ATTAATAATACTGAAATTAATTAATAATGAGCTACTAATGAAGAAAAAGTTACAGATA 3364

OY      1034  AGGATGAAATTAAGTGTGAATTTGACGATTCATTAATAGTAGAAGAAAGATGATGAG 1093
      |||||
DB      3363  TTGACACAGTTGAGAGAGCATATGAAGAGTTTAAATATTAATTAAGCAGCAACTCTA 3304

OY      1094  ACTTAGACGAATTTAAGCTTATGTCATATGACATATGACATTTCCAAAGATGAAGAAAACTAG 1153
      |||||
DB      3303  CAGGTGATTAAGTACTGCTTAAGATACAGCACTACAAAAAGTTCAACAACTTCATCGAA 3244

OY      1154  GAATTTATTAAGAACTAGAAAGATTGATAGAGAAAAATGAAAAATTTAGATGATTTAGATG 1213
      |||||
DB      3243  ATCTGTTTAAGAAACACAGAGGTAATAAAGAAATTAATCATCAAGCTGAGTAAAGAAA 3184

OY      1214  AAGGATGAAAAATATATAGAAAGATTTCTGAAGAAAAATTAATAAAGAAAAAGAAAT 1273
      |||||
DB      3183  CCAAAATGAAACAAACACCAAAATG---CATCAACAAGAAATTAATGATGACAAAACAG 3127

OY      1274  ATGAAAAAAACAAAGATATAATATTTTAAACCAATG 1309
      |||||
DB      3126  AAGTTGATCTGATTAATTAATCAAGCAAAACAAATG 3091

RESULT 13
US-11-082-746-2/c
; Sequence 2, Application US/11082746
; Publication NO. US20050244861A1
; GENERAL INFORMATION:
; APPLICANT: Sundetrom, Paula
; TITLE OF INVENTION: NUCLEIC ACIDS ESSENTIAL FOR EXPRESSION OF HYPLAL-SPECIFIC GENES
; TITLE OF INVENTION: AND METHODS FOR USING THE SAME
; FILE REFERENCE: DC-0300
; CURRENT APPLICATION NUMBER: US/11/082,746

```

```

CURRENT FILING DATE: 2005-03-17
PRIORITY APPLICATION NUMBER: US 09/725,010
PRIOR FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: US 60/167,672
PRIOR FILING DATE: 1999-11-29
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.3
SEQ ID NO 2
LENGTH: 1434
TYPE: DNA
ORGANISM: Candida albicans
US-11-082-746-2

Query Match      4.3%; Score 63.2; DB 9; Length 1434;
Best Local Similarity 52.6%; Pred. No. 0.37;
Matches 160; Conservative 0; Mismatches 143; Indels 1; Gaps 1;

OY    1107 TAACCCATTTGTGCATATGCAATTTCCTCAGATGAAGAAAACCTAGAAATTATAAGA   1166
Db     305 TTTAAACCACGTAATTAAGCTTAGAACTTAACTATGTGTAAAAAAGGGGTATTATTAAAA   246

OY    1167 ACTAGAAGATTGTGTAGAGAAAATGAAAATTTAGATGATTGATGAAGAAATAGAAAA   1226
Db     245 GCTATCAACTATGAAACCGAAAACGAAAACAAACAAGTAGAACAGAAAAGTTAA   186

OY    1227 ATCATCAGAAAGATTAATCTGGAAGAAAATTTAAAAAAGAAAAGAAATATGAAAAACAA   1286
Db     185 ACAATTTGAAACAAAACAATTTTTTAAAAACCTAAAAGTACCTATCACCTTTAATGTATGTA   126

OY    1287 GGATTAATATTTTAAACCAAATGATTAAGTTTG-TATGATGACATATTTAAAAATATA   1345
Db     125 CAAAACAATTTCAATTAATTAATGACTCTTGCCAATAATATTGAAAACCCAAATGTA   66

OY    1346 AAATGATTAACAGGTTAATAAGAAAAGAAAATTCATTAATCATTTGTCATATAT   1405
Db     65 ACAATCAAAATTAGAAAAAAGAAATCCCAAAAAAGAACAAACTTAGATTCAGTTAT   6

OY    1406 TTGA 1409
Db     5 TAGA 2

RESULT 14
US-10-909-125-825
Sequence 825, Application US/10909125
Publication No. US2005026121B1
GENERAL INFORMATION:
APPLICANT: Esau, Christine
APPLICANT: Lollo, Bridget
APPLICANT: Bennett, C. Frank
APPLICANT: Freiler, Susan M.
APPLICANT: Grilley, Richard H.
APPLICANT: Baker, Brenda F.
APPLICANT: Vickers, Timothy
APPLICANT: Marcuseon, Eric G.
APPLICANT: Koller, Erich
APPLICANT: Swayze, Eric
APPLICANT: Jain, Ravi
APPLICANT: Bhac, Balkrishen
APPLICANT: Petalite, Egen
TITLE OF INVENTION: Oligomeric Compounds And Compositions For Use In Modulation
FILE REFERENCE: ISIS0080-100 (CORE0016US)
CURRENT APPLICATION NUMBER: US/10/909,125
CURRENT FILING DATE: 2004-07-30
PRIOR APPLICATION NUMBER: US 60/492,056
PRIOR FILING DATE: 2003-07-31
PRIOR APPLICATION NUMBER: US 60/516,303
PRIOR FILING DATE: 2003-10-31
PRIOR APPLICATION NUMBER: US 60/531,596
PRIOR FILING DATE: 2003-12-19
PRIOR APPLICATION NUMBER: US 60/562,417
PRIOR FILING DATE: 2004-04-14

```

	NUMBER OF SEQ ID NOS: 2184	
	SOFTWARE: FastSeq for Windows Version 4.0	
	SEQ ID NO 825	
	LENGTH: 6189	
	TYPE: DNA	
	ORGANISM: H. sapiens	
	US-10-909-125-825	
	Query Match	4.3%; Score 63.2; DB 6; Length 6189;
	Best Local Similarity	42.2%; Pred. No. 0.37;
	Matches 517; Conservative	0; Mismatches 688; Indels 21; Gaps 2;
QY	133	AGACTTGGCTAAAGAAAAGTTGCTCAAGAAACAACAAGCATCTTAGACACAAGAGACGTCGT 192
DB	1894	ACATTTTGGAAAATTGAGATTCAGATTCGACACTGAAAAAATGATTAAGAAAGTTAAACCTTTAA 1953
QY	193	AAAGAAAAGTTGCAGAACAACAAGGCGATTTAGACAAGAGAGACGTCGTTAAAGAAAAG 252
DB	1954	AAACATTAATTGAAAGCTTAAGAGAGGAAAAAGAAAAGCATTAATAAAGAAATTTGAAGCT 2013
QY	253	TTGCAAGAACCAACAAGCGATTTTAGAA CAAGATGACTTCTTAAGAAAAGTTACAAAG 312
DB	2014	GAAGAAGAAAATAACAAACTAGAGATAGTCCAAAGAACTCGAGAGAGAGTGCGAATTT 2073
QY	313	CAGCAAAAGCATTTAGAACCAAGAGACGTCCTAAAGAAAAGTTGCAAGAACCAAGC 372
DB	2074	GATAGAGAAATTTTGGAAAAGAAATTTTAAAAAGTATGAACTGAAGATCTCTTTT 2133
QY	373	GATTTAGAACCAAGAGACGTCCTAAAGAAAAGTTGCAAGAACCAAGCGATTTAGAA 432
DB	2134	AATATGGACACTGAAATCCTTAACATTGAAAAAATCAAAATTTGGAAAAAATCATCAA 2193
QY	433	CAAGAGAGACTTGGCTAAAGAAAAGTTGCAAGAACCAAGAAAGCGATTTAGAACCAAGAGAG 492
DB	2194	GATGATTAATGAACCAAGAGAAAACATGTCGTCAAAAGAGAGAACTTTAAAGAGAAAGCA 2253
QY	493	CGTCGTAAGAAAAGTTGCAAGAACCAAGAAAGCATTTAGAACCAAGAGAGACGTCCTAA 552
DB	2254	GACAAAGATTTAAAAAGAAAAGCGAAGAAATCTTTTAGGAGGAAAAAATAAAGATCTTAA 2313
QY	553	GAAGAAGTTGCAAGAACCAAGAGGATTTTAGAACCAAGAGACGTCGTTAAAGAAAGTTG 612
DB	2314	GAAGAGAGAAAACATACCAAGATTAAGACTCAAGAAATTTACTCTTTGGGTATGAGT 2373
QY	613	CAAGAGACGCA-----AGAGATTTTAGAACCAAGAAAGCGTCGTTAGCAAAAAT 663
DB	2374	GCCATGAGGAATCTATAGGGCTCATTTAGTGGAAAAAGAAATAGACATTGAAAAACAA 2433
QY	664	TTAGAAAAGAAAAGAACATGAGATATATTAGCAAGGATTTATATGTCGTTTAGAA 723
DB	2434	GAAGAAGCATTAAGAAAGAAAGTAAAGAAAACCTGAGAAGCGATCTCAAAATTAAGAAAAG 2493
QY	724	ATACAGGATATGAACCTTCATCGAAAAATGACGTCGATTTTATATCCATCAATCT 783
DB	2494	GACATTTGGAAGATGAGAAAAGAAAACCTTTGAAAAAAGAAAAGAAATTAACATGACAT 2553
QY	784	TCTTTACCTCGAGCAACAGAGGGAATAGTAGAATTCACAAAGAAATATCTATATAGAA 843
DB	2554	AAATCAGAAAAAGCAAAATTTGATCTTAGTGAATGTGTTGATTAATAAAGAAAGAAAGAC 2613
QY	844	AAAACAATAGAGATCTTATTAACAACAATGTTAGAGACGAAGGGAATATCATTAAGGA 903
DB	2614	AAAGTATATTCGACTCACACAGAAAATATGCATTAAGAAAGGTGAGAGAGCAAAAAATATCT 2673
QY	904	CATCTTGAGAAAAGAAAGATGCTTCAATTAACCAAGAACAAAAGAAAGATTAATCTGCT 963
DB	2674	GCTGCTATTAATAAAAACTGACGAGAGAGAAAAGTAGAAGAAAAGATGAGTAAGAAACAT 2733
QY	964	GACATTAACAAATCACTACCTTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCAATA 1023
DB	2734	GAACAAAGAAAAGCCTGAAAGAAAGAA-----GGCATCTAGCAAGAAAGCAAAAGAA 2781
QY	1024	AGTAAGTATGAGATGAATTAAGTCTGAAATATGACGATTCATTAATAGATGAAGAAAG 1083

```
Db      2782  AACGCTTGATGAGAAAAAATAAACAATCAGATTAATAGTAAATACGTAATCAGAA 2841
Qy      1084  GATGATGAAGACTTATGACCAATTTAAAGCTTATGCAATATGACAAATTTCCAAAGATGA 1143
Db      2842  AAGGCAAAAAATGAAGAAAAAGACAGGAGCTAGATTAAGAAAAATCTTAGAATATAA 2901
Qy      1144  GAAAACTATGAAATTTATTAAGAACTAGAAATTTGATGAGAAAAAATGAAATTTAGAT 1203
Db      2902  GAAAGTATTAATTAATTAATCTACTCCAAACCATACAGAGAAAGAAAAATCAAGTATAGTA 2961
Qy      1204  GATTAGATGAAGAAATAGAAAAATCATCAGAAATTAATCTGAGAAAAAATTAATAAAAA 1263
Db      2962  GACGGTAATTAAGCAACAATGAAAAACCTTATCCCTTAAGAAAAAACAAGATGAA 3021
Qy      1264  GGAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTTGTAT 1323
Db      3022  CCTTGAAAACTCCAGATGAAAAAGAAAAAGATTAAGATTAATAGATGATAC 3081
Qy      1324  GATGACATATTTAAAAATATAAAAA 1349
Db      3082  AAGAACGACGACAAACATTAAGATTA 3107
```

```
RESULT 15
US-10-793-626-693
; Sequence 693, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 693
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-693
```

```
Query Match      4.2%; Score 62.8; DB 6; Length 627;
Best Local Similarity 48.6%; Pred. No. 0.41;
Matches 203; Conservative 0; Mismatches 212; Indels 3; Gaps 1;
```

```
Qy      1061  ATTCATTATATGATGAAGAAAGATGATGAAGACTTAAAGCAATTTAAAGCTATGTGC 1120
Db      134  ATCAAGTAAGTAAAAAATTAACGAATTAAGAAAGAAAAAGCAAAAAATTTATTTAAAAAG 193
Qy      1121  AATATGACAATTTCCAAAGATGAAGAAAAACATAGCAATTTATTAAGAACTAGAAATTTGA 1180
Db      194  TAAATGATTAAGTCAAGACACGTAAGAAAAACAGCTGAGATATAGTTGAAAAATGTAA 253
Qy      1181  TAGAGAAAAATGAAATTTAGATTTAGATGAGAAATAGAAAAATCATCAGAAAGAT 1240
Db      254  AACAAAGACAAAAGAAATTTGAAAAAGAAAGAGCTAGATTAATCTGAAAAAGCAT 313
Qy      1241  TATCTGAAGAAAAATTAAGAAAAAGAAAGATATGAAAAAACAAGATTAATTTTA 1300
Db      314  TTAACAGACCAAGCATATCTTGAAACATGTAGAAAAACAAGCAAGAAAAAGAAAGTTG 373
Qy      1301  AACCAAAATGATAAAGTTGTATGATGACATATTTAAAAAATTAATAATGATAAGCAG 1360
Db      374  AACCACTTATATGTCTAT--TAAAGAAAAATATATAATCAATGATGCTTATGCAAG 430
Qy      1361  TTAATTAAGAAAAAGAAAAATTCATTAATTCATTTGTTTCAATATATTGACGAGACATG 1420
```

```
Db      431  CTTACAAAAAGCACTTAATTAAGAAAAAGAACTGTTTTCTTATTTGAATGAAGATAATG 490
Qy      1421  AAATTTACAGATGCGATGATGATTATCTGAAGATATTAATTAATTTATGAAACT 1478
Db      491  CACACAAATCGAAATAGACGAAAAATGAAAGATCTTTCTTAAGCATATTAAGAAAT 548
```

```
Search completed: December 7, 2005, 06:43:40
Job time : 598 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 01:48:14 ; Search time 1329 Seconds
(without alignments)
9221.383 Million cell updates/sec

Title: us-09-837-344-41

Perfect score: 1482

Sequence: 1 CAGAGACACACAGGATCT.....ATATTTTATGAACTATA 1482

Scoring table: IDENTITY NUC

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA Main.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1482	100.0	1482	3	US-09-837-344-41
2	1482	100.0	1482	3	US-09-900-963-41
3	1480.4	99.9	1482	3	US-09-837-344-45
4	1480.4	99.9	1482	3	US-09-900-963-45
5	1446	97.6	1493	3	US-09-837-344-38
6	1446	97.6	1493	3	US-09-900-963-38
7	950	64.1	950	3	US-09-837-344-32
8	950	64.1	950	3	US-09-900-963-32
9	472.8	31.9	988	3	US-09-837-344-34
10	472.8	31.9	988	3	US-09-900-963-34
11	470.8	31.8	954	3	US-09-837-344-37
12	470.8	31.8	954	3	US-09-900-963-37
13	233.4	15.7	1300	7	US-10-668-749A-1
14	153.2	10.2	1297	8	US-10-425-115-67687
15	151.2	10.2	6292	7	US-10-221-714A-661
16	149.2	10.1	6767	7	US-10-221-714A-330
17	146	9.9	3662	10	US-11-097-143-5152
18	145.8	9.8	1243	8	US-10-425-115-172717
19	145	9.8	1204	8	US-10-437-963-77858
20	142.4	9.6	1081	8	US-10-425-115-16756
21	139.4	9.4	9539	6	US-10-239-676-52
22	139.4	9.4	9539	6	US-10-240-453-54
23	137.8	9.3	6668	6	US-10-311-455-1670

24	132.8	9.0	778	8	US-10-363-345A-2179	Sequence 2179, Ap
25	132.8	9.0	778	8	US-10-363-345A-2180	Sequence 2180, Ap
26	132.8	9.0	778	9	US-10-363-483A-2179	Sequence 2179, Ap
27	132.8	9.0	778	9	US-10-363-483A-2180	Sequence 2180, Ap
28	131.4	8.9	7442	7	US-10-221-714A-409	Sequence 409, App
29	129.8	8.8	921	8	US-10-425-115-38710	Sequence 38710, A
30	129.8	8.8	1762	8	US-10-363-345A-7529	Sequence 7529, Ap
31	129.8	8.8	1762	8	US-10-363-345A-7530	Sequence 7530, Ap
32	129.8	8.8	1762	9	US-10-363-483A-7529	Sequence 7529, Ap
33	129.8	8.8	1762	9	US-10-363-483A-7530	Sequence 7530, Ap
34	129.2	8.7	1995	8	US-10-363-345A-7533	Sequence 7533, Ap
35	129.2	8.7	1995	8	US-10-363-345A-7534	Sequence 7534, Ap
36	129.2	8.7	1995	9	US-10-363-483A-7533	Sequence 7533, Ap
37	129.2	8.7	1995	9	US-10-363-483A-7534	Sequence 7534, Ap
38	129	8.7	1200	8	US-10-363-345A-25801	Sequence 25801, A
39	129	8.7	1200	8	US-10-363-345A-25802	Sequence 25802, A
40	129	8.7	1200	9	US-10-363-483A-25801	Sequence 25801, A
41	129	8.7	1200	9	US-10-363-483A-25802	Sequence 25802, A
42	129	8.7	7597	6	US-10-311-455-986	Sequence 986, App
43	129	8.7	14006	6	US-10-311-455-1931	Sequence 1931, Ap
44	128.6	8.7	3673778	6	US-10-312-841-1	Sequence 1, Appl
45	128.6	8.7	3673778	6	US-10-312-841-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-837-344-41
Sequence 41, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ. ID NO.: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

```
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-837-344-41

Query Match      100.0%; Score 1482; DB 3; Length 1482;
Best Local Similarity 100.0%; Pred. No. 2,4e-196;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGAACACAAAGCGATCTAGAACAAAGAGACGCTTAAAGAAAGTTGCAAGACAA 60
DB 1 CAAGAACACAAAGCGATCTAGAACAAAGAGACGCTTAAAGAAAGTTGCAAGACAA 60
QY 61 CAAGCGATTTAGAACAAAGATAGACTTGTAAAGAAAGTTACAAAGACGAAAGCGAT 120
DB 61 CAAGCGATTTAGAACAAAGATAGACTTGTAAAGAAAGTTACAAAGACGAAAGCGAT 120
QY 121 TTAGAACAAAGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGACAA 180
DB 121 TTAGAACAAAGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGACAA 180
QY 181 GAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGT 240
DB 181 GAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGT 240
QY 241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTCTAAAGAA 300
DB 241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTCTAAAGAA 300
QY 301 AAGTTACAAAGAGACGAAAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAGTTGCAA 360
DB 301 AAGTTACAAAGAGACGAAAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAGTTGCAA 360
QY 361 GAAACAAAGAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAGTTGCAAGACAA 420
DB 361 GAAACAAAGAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAGTTGCAAGACAA 420
QY 421 AGCGATTTAGAACAAAGAGACCTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 480
DB 421 AGCGATTTAGAACAAAGAGACCTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 480
QY 481 GAAACAAAGAGACGCTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGACAAAG 540
DB 481 GAAACAAAGAGACGCTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGACAAAG 540
QY 541 AGACGCTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTGCT 600
DB 541 AGACGCTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTGCT 600
QY 601 AAAGAAAGAGTTGCAAGAGACGAAAGATTTAGAACAAAGAGAGCTGATTCGAAAAA 660
DB 601 AAAGAAAGAGTTGCAAGAGACGAAAGATTTAGAACAAAGAGAGCTGATTCGAAAAA 660
QY 661 AATTAGAAAGAAAGAAAGCACTGAGATATATTGACAGAGATTTATATGCTCGTTTA 720
DB 661 AATTAGAAAGAAAGAAAGCACTGAGATATATTGACAGAGATTTATATGCTCGTTTA 720
QY 721 GAAATCCGCTTAACTTCACTTCATCAGAAAAAGAAAGTGAATATATACACACTCAA 780
DB 721 GAAATCCGCTTAACTTCACTTCATCAGAAAAAGAAAGTGAATATATACACACTCAA 780
QY 781 TCTTCTTACCTCAGACAAACAGAGGAAATAGAGATTCACAGGAAATATCTATATA 840
DB 781 TCTTCTTACCTCAGACAAACAGAGGAAATAGAGATTCACAGGAAATATCTATATA 840
QY 841 GAAAAAACAATAGAGATCTATTACAAACAATGTTGAAGAGCAAGGATATACATATA 900
DB 841 GAAAAAACAATAGAGATCTATTACAAACAATGTTGAAGAGCAAGGATATACATATA 900
```

```
QY 901 GGACATCTTGAAGAAAGAAAGATGCTTCAATTAACACAGAACAAAGAAAGATTAATCT 960
DB 901 GGACATCTTGAAGAAAGAAAGATGCTTCAATTAACACAGAACAAAGAAAGATTAATCT 960
QY 961 GCTGACATACAAATCATATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCAA 1020
DB 961 GCTGACATACAAATCATATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCAA 1020
QY 1021 ATTAAGTATGATGAGATGAAGAAATAGCTGAAATATGACGATTCCTTAATAGATGAAGAA 1080
DB 1021 ATTAAGTATGATGAGATGAAGAAATAGCTGAAATATGACGATTCCTTAATAGATGAAGAA 1080
QY 1081 GAAGATGATGAAGACTTGAAGCAATTTAAGCTATATGCAATATGACAAATTTCCAGAT 1140
DB 1081 GAAGATGATGAAGACTTGAAGCAATTTAAGCTATATGCAATATGACAAATTTCCAGAT 1140
QY 1141 GAAGAAACATAGCAATTTATTAAGAACTAGAGATTTGATAGAGAAAAATGAATAATTTA 1200
DB 1141 GAAGAAACATAGCAATTTATTAAGAACTAGAGATTTGATAGAGAAAAATGAATAATTTA 1200
QY 1201 GATGATTTAGATGAAGCAATGAAGAAATCATCAGAGAAATTTATCGAAGAAAAATTAATA 1260
DB 1201 GATGATTTAGATGAAGCAATGAAGAAATCATCAGAGAAATTTATCGAAGAAAAATTAATA 1260
QY 1261 AAAGAAAGAAATATGAAGAAACAAAGGATTAATTTTAAACCAATGATTAAGATTG 1320
DB 1261 AAAGAAAGAAATATGAAGAAACAAAGGATTAATTTTAAACCAATGATTAAGATTG 1320
QY 1321 TATGATGACATATTAATAATTAATAATGATTAAGCAGTTAATAGAAAGAAAGAA 1380
DB 1321 TATGATGACATATTAATAATTAATAATGATTAAGCAGTTAATAGAAAGAAAGAA 1380
QY 1381 TTCATTAATATCATGTTTCATATATTTGACGAGACATAGAAATTTACAGATCGTGAT 1440
DB 1381 TTCATTAATATCATGTTTCATATATTTGACGAGACATAGAAATTTACAGATCGTGAT 1440
QY 1441 GAGTATCTGAAGATTAATTAATTAATTTTGAACACTATA 1482
DB 1441 GAGTATCTGAAGATTAATTAATTAATTTTGAACACTATA 1482

RESULT 2
US-09-900-963-41
; Sequence 41, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUG/ME, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; OF STIMULATING THE T LYMPHOCYTES
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
```


COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-837-344-45

Query Match 99.9%; Score 1480.4; DB 3; Length 1482;
Best Local Similarity 99.9%; Pred. No. 4e-196;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CAAGAACCAAAAGCATCTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCA 60
1 CAAGAACCAAAAGCATCTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCA 60
61 CAAGGATTTAGAACCAAGATAGACTTGTAAAGAAAAGTTGCAAGAGACGAT 120
61 CAAGGATTTAGAACCAAGATAGACTTGTAAAGAAAAGTTGCAAGAGACGAT 120
121 TTAGAACCAAGAGACCTTCTAAAGAAAAGTTGCAAGAACCAAGAGACGAT 180
121 TTAGAACCAAGAGACCTTCTAAAGAAAAGTTGCAAGAACCAAGAGACGAT 180
181 GAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGCTTAAGAAAAGTTGCA 240
181 GAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGCTTAAGAAAAGTTGCA 240
241 GCTAAAGAAAAGTTGCAAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCA 300
241 GCTAAAGAAAAGTTGCAAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCA 300
301 AAGTTACCAAGAGACGATTTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAG 360
301 AAGTTACCAAGAGACGATTTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAG 360
361 GAAACAACAAAGCATTTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCA 420
361 GAAACAACAAAGCATTTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCA 420
421 AGGATTTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGAT 480
421 AGGATTTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGAT 480
481 GAAACAAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGATTTAGAACCA 540
481 GAAACAAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGATTTAGAACCA 540

481 GAAACAAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGATTTAGAACCA 540
541 AGGATTTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGCT 600
541 AGGATTTAGAACCAAGAGACGCTTAAGAAAAGTTGCAAGAACCAAGAGACGCT 600
601 AAAGAAAAGTTGCAAGAGACCAAGAGATTTAGAACCAAGAGAGCTGATTCGAAA 660
601 AAAGAAAAGTTGCAAGAGACCAAGAGATTTAGAACCAAGAGAGCTGATTCGAAA 660
661 AATTAGAAAAGAAAAGCAATGAGATTTATTTAGAGAGAGATTTATTTAGTCTT 720
661 AATTAGAAAAGAAAAGCAATGAGATTTATTTAGAGAGAGATTTATTTAGTCTT 720
721 GAAATACCAAGCTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 780
721 GAAATACCAAGCTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 780
781 TCTTCTTTACCTCAGAGCAAGAGAGATTTAGAGATTTCCAGAAAATTTATTA 840
781 TCTTCTTTACCTCAGAGCAAGAGAGATTTAGAGATTTCCAGAAAATTTATTA 840
841 GAAAAAACAATTAAGATTTATTTACCAAAATTTGCAAGAGAGATTTATTA 900
841 GAAAAAACAATTAAGATTTATTTACCAAAATTTGCAAGAGAGATTTATTA 900
901 GAAATACCAAGCTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 960
901 GAAATACCAAGCTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 960
961 GCTGACATACCAAAATCATACATTTAGAGACAGTAAATTTCTGATTTATGATTT 1020
961 GCTGACATACCAAAATCATACATTTAGAGACAGTAAATTTCTGATTTATGATTT 1020
1021 AATAAGTATGATAGAGATTAAGAAATTAAGTCTGATTTATGATTTATTA 1080
1021 AATAAGTATGATAGAGATTAAGAAATTAAGTCTGATTTATGATTTATTA 1080
1081 GAAATACCAAGCTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 1140
1081 GAAATACCAAGCTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 1140
1141 GAAAGAAAATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 1200
1141 GAAAGAAAATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 1200
1201 GATGATTTAGTGAAGAAATTAAGAACTTCCATCAGAAAATGAGATTTATTT 1260
1201 GATGATTTAGTGAAGAAATTAAGAACTTCCATCAGAAAATGAGATTTATTT 1260
1261 AAAGAAAAGTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 1320
1261 AAAGAAAAGTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 1320
1321 TATGATGACATTTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 1380
1321 TATGATGACATTTATTAAGAACTTCCATCAGAAAATGAGATTTATTTATACCAT 1380
1381 TTGATTAATATCAATTTGATTAATTTAGAGAGATTTATTTATGATTTAT 1440
1381 TTGATTAATATCAATTTGATTAATTTAGAGAGATTTATTTATGATTTAT 1440
1441 GAGTATCTGAAGATTAATTAATTTATTTATTTATTTATTTATTTATTTAT 1482
1441 GAGTATCTGAAGATTAATTAATTTATTTATTTATTTATTTATTTATTTAT 1482

RESULT 4
US-09-900-963-45
Sequence 45, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUBERIN-MARCHAND, Claudine

DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/900,963
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-900-963-45
Query Match 99.9%; Score 1480.4; DB 3; Length 1482;
Best Local Similarity 99.9%; Pred. No. 4e-196;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 301 AAGTACAAAGACGCAAAAGCATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAA 360
Qy 361 GAACAAACAAAGCATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAA 420
Db 361 GAACAAACAAAGCATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAA 420
Qy 421 AGCGATTTAGAACAAAGAGACCTTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
Db 421 AGCGATTTAGAACAAAGAGACCTTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
Qy 481 GAACAAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAAGCATTTAGAACAAAG 540
Db 481 GAACAAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAAGCATTTAGAACAAAG 540
Qy 541 AGACGCTCTAAAGAAAAGTTGCAAGAACAAAGCATTTAGAACAAAGAGACGCTCT 600
Db 541 AGACGCTCTAAAGAAAAGTTGCAAGAACAAAGCATTTAGAACAAAGAGACGCTCT 600
Qy 601 AAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGCCTGATACGAAAAA 660
Db 601 AAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGCCTGATACGAAAAA 660
Qy 661 AATTTAGAAAAGAAAAGAACATGAGATATATTAGCAGAGATTTATATGCTGTTTA 720
Db 661 AATTTAGAAAAGAAAAGAACATGAGATATATTAGCAGAGATTTATATGCTGTTTA 720
Qy 721 GAAATACACGCTATAGAACTTCCATCAGAAAATGAAACGCTGATTTATACCAATCAA 780
Db 721 GAAATACACGCTATAGAACTTCCATCAGAAAATGAAACGCTGATTTATACCAATCAA 780
Qy 781 TCTTCTTTTACTCAGGCAACAGAGGGAATGATGAGATTCCAAAGGAAATATCTATATA 840
Db 781 TCTTCTTTTACTCAGGCAACAGAGGGAATGATGAGATTCCAAAGGAAATATCTATATA 840
Qy 841 GAAAAAAACAAATAGAGATCTATTACAAACAAATGTTGAAGAGCAAGAGATATACATATA 900
Db 841 GAAAAAAACAAATAGAGATCTATTACAAACAAATGTTGAAGAGCAAGAGATATACATATA 900
Qy 901 GGAATCTTTGAAGAAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAATTAATCT 960
Db 901 GGAATCTTTGAAGAAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAATTAATCT 960
Qy 961 GCTGACATTAACAATATCATTAATGAGACAGTAAATTTCTGATGTTAATGATTTCAA 1020
Db 961 GCTGACATTAACAATATCATTAATGAGACAGTAAATTTCTGATGTTAATGATTTCAA 1020
Qy 1021 ATAAGTATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1080
Db 1021 ATAAGTATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1080
Qy 1081 GAAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1140
Db 1081 GAAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1140
Qy 1141 GAAGAAAACATAGAAATTTTAAAGAACTGAGAAATTTGATAGAGAAAATGAAAATTTA 1200
Db 1141 GAAGAAAACATAGAAATTTTAAAGAACTGAGAAATTTGATAGAGAAAATGAAAATTTA 1200
Qy 1201 GATGATTTAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1260
Db 1201 GATGATTTAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1260
Qy 1261 AAGGAAAAGAAATATGAGAAAACAAAGATTAATTTTAAACCAATGATTAAGTTTG 1320
Db 1261 AAGGAAAAGAAATATGAGAAAACAAAGATTAATTTTAAACCAATGATTAAGTTTG 1320
Qy 1321 TATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1380
Db 1321 TATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1380
Qy 1381 TTGATTAATATCATTTGTTTCAATATATTGAGAGAGACAAATTAATTTTACAGATGCTGAT 1440
Db 1381 TTGATTAATATCATTTGTTTCAATATATTGAGAGAGACAAATTAATTTTACAGATGCTGAT 1440

QY 1441 GAGTATCTGAAGATTAATAATTTTATGAACATAA 1482
DB 1441 GAGTATCTGAAGATTAATAATAATTTTATGAACATAA 1482

RESULT 5
US-09-837-344-38
Sequence 38, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierrie
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STRGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 38:

US-09-837-344-38

Query Match 97.6%; Score 1446; DB 3; Length 1493;
Best Local Similarity 99.8%; Pred. No. 2.4e-191;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1 CAAGAACCAAGCATCTTAGAACAGAGACGCTGAAGAAAGTTGCAAGACAA 60
DB 1 CAAGAACCAAGCATCTTAGAACAGAGACGCTGAAGAAAGTTGCAAGACAA 60
QY 61 CAAGGATTTTGAAGACATGACTTGCTAAAGAAAGTTACAAAGCAGCAAGCGAT 120
DB 61 CAAGGATTTTGAAGACATGACTTGCTAAAGAAAGTTACAAAGCAGCAAGCGAT 120
QY 121 TTGAACAAGAGACTTGCTTAAGAAAGTTGCAAGAACCAAGGCGATCTAGAACAA 180
DB 121 TTGAACAAGAGACTTGCT- AAAGAAAGTTGCAAGAACCAAGGCGATCTAGAACAA 179

QY 181 GAGAGACGTGCTAAAGAAAGTTGCAAGACAAAGCGATTTAGAACAGAGAGCGT 240
DB 180 GAGAGACGTGCTAAAGAAAGTTGCAAGACAAAGCGATTTAGAACAGAGAGCGT 239
QY 241 GCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGATGACTTGCTAAAGAA 300
DB 240 GCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGATGACTTGCTAAAGAA 299
QY 301 AAGTTACAAAGACGCAAGCGATTTAGAACAGAGACGCTGCTAAAGAAAGTTGCA 360
DB 300 AAGTTACAAAGACGCAAGCGATTTAGAACAGAGACGCTGCTAAAGAAAGTTGCA 359
QY 361 GAACAAACAAGCGATTTAGAACAGAGACGCTGCTAAAGAAAGTTGCAAGACAA 420
DB 360 GAACAAACAAGCGATTTAGAACAGAGACGCTGCT- AAGAAAGTTGCAAGACAA 418
QY 421 AGCGATTTAGAACAGAGACGCTGCTAAAGAAAGTTGCAAGAACCAAGCGATTTA 480
DB 419 AGCGATTTAGAACAGAGACGCTGCTAAAGAAAGTTGCAAGAACCAAGCGATTTA 478
QY 481 GAACAGAGACGCTGCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGAG 540
DB 479 GAACAGAGACGCTGCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGAG 538
QY 541 AGAGCTGCTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGAGAGAGCGTGT 600
DB 539 AGAGCTGCT- AAGAAAGTTGCAAGAACCAAGCGATTTAGAACAGAGAGAGCGTGT 597
QY 601 AAAGAAAGTTGCAAGACAGCAAGCAAGATTTAGAACAAAGAGCGTATGCAAAAA 660
DB 598 AAAGAAAGTTGCAAGACAGCAAGCAAGATTTAGAACAAAGAGCGTATGCAAAAA 657
QY 661 AATTAGAAAGAAAGAAAGCAATGAGATATTTAGCAGAGATTTATATGCTGTTA 720
DB 658 AATTAGAAAGAAAGAAAGCAATGAGATATTTAGCAGAGATTTATATGCTGTTA 717
QY 721 GAATATCCAGCATATGAACTTCCATCAGAAATGAAAGCGATATTTATATACCATCA 780
DB 718 GAATATCCAGCATATGAACTTCCATCAGAAATGAAAGCGATATTTATATACCATCA 777
QY 781 TCTTCTTTACCTCAGAGCAACAGAGGAAATGTAAGATTCCAGAAATATCTATATA 840
DB 778 TCTTCTTTACCTCAGAGCAACAGAGGAAATGTAAGATTCCAGAAATATCTATATA 837
QY 841 GAAAAAACAATAGGAATCTATTCAACAAATGTTGAAGACGAAGGATTTACATAA 900
DB 838 GAAAAAACAATAGGAATCTATTCAACAAATGTTGAAGACGAAGGATTTACATAA 897
QY 901 GGACATCTTGAAGAAAGAAAGATGCTTCAATPAAACGAAACAAAGAAATGATTAATCT 960
DB 898 GGACATCTTGAAGAAAGAAAGATGCTTCAATPAAACGAAACAAAGAAATGATTAATCT 957
QY 961 GCTGACATACAAATCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTCA 1020
DB 958 GCTGACATACAAATCATACATTAGAGACAGTAAATATTTCTGATGTTAATGATTTCA 1017
QY 1021 AATAAGTATGATGAGATGAATTAAGTGTGTAATGACGATTCATTAATAGTAAAGAA 1080
DB 1018 AATAAGTATGATGAGATGAATTAAGTGTGTAATGACGATTCATTAATAGTAAAGAA 1077
QY 1081 GAAGATGATGAAGACTTGAAGCAATTTAAGCTATTTGCAATPAGACAAATTTCCAAAGAT 1140
DB 1078 GAAGATGATGAAGACTTGAAGCAATTTAAGCTATTTGCAATPAGACAAATTTCCAAAGAT 1137
QY 1141 GAAGAAACATAGAAATTTATPAAAGACATAGAAATTTGATAGAGAAAAATGAAATTTTA 1200
DB 1138 GAAGAAACATAGAAATTTATPAAAGACATAGAAATTTGATAGAGAAAAATGAAATTTTA 1197
QY 1201 GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGATTTCTGAAGAAAAATPAAA 1260
DB 1198 GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGATTTCTGAAGAAAAATPAAA 1257

1141 GAAGAAAACATGGAATTTATTAAGAACTAGAGATTTGATGAGAAAAATGAATAATTTA 1200
1138 GAAGAAAACATGGAATTTATTAAGAACTAGAGATTTGATGAGAAAAATGAATAATTTA 1197
1201 GATGATTTGATGAGAAATAGAAAAATCATCGAAGAAATTTATCTGAGAAAAATTAATA 1260
1198 GATGATTTGATGAGAAATAGAAAAATCATCGAAGAAATTTATCTGAGAAAAATTAATA 1257
1261 AAAGAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTTTG 1320
1258 AAAGAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTTTG 1317
1321 TATGATGAGCAATTTTAAATAATTAATAATGATTAAGCAATTAAGAAAAA 1380
1318 TATGATGAGCAATTTTAAATAATTAATAATGATTAAGCAATTAAGAAAAA 1377
1381 TTCATTAATCATTTGTTTATATATTTGACGAGCAATGAATTTTACAGATCGTGAT 1440
1378 TTCATTAATCATTTGTTTATATATTTGACGAGCAATGAATTTTACAGATCGTGAT 1437
1441 GAGTATCTGAAGATTTAACTAAATTTTATGAAACTATTA 1482
1438 GAGTATCTGAAGATTTAACTAAATTTTATGAAACTATTA 1479

RESULT 7
US-09-837-344-32
Sequence 32, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-837-344-32
Query Match 64.1%; Score 950; DB 3; Length 950;
Best Local Similarity 100.0%; Pred. No. 1,2e-122;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

11 AAAGCATCTGAAACAAAGAGACGCTGCTAAAGAAAACCTCCAGAACAAACACCGATT 70
1 AAAGCATCTGAAACAAAGAGACGCTGCTAAAGAAAACCTCCAGAACAAACACCGATT 60
71 TAGAACATGATGACTGCTTAAAGAAAAGTTTAAAGAGACGCAAGACGATTAGAACAG 130
61 TAGAACATGATGACTGCTTAAAGAAAAGTTTAAAGAGACGCAAGACGATTAGAACAG 120
121 AGAGACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGACGCTG 180
121 AGAGACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGACGCTG 190
191 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGACGCTTAAAGAAA 250
181 CTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGACGCTTAAAGAAA 240
251 AGTTCAGAGAACAAAGAGCGATTTTGAACAGAGACGCTTAAAGAAAAGTTTACAAG 310
241 AGTTCAGAGAACAAAGAGCGATTTTGAACAGAGACGCTTAAAGAAAAGTTTACAAG 300
311 AGCAGCAAGCGATTTTGAACAGAGACGCTTAAAGAAAAGTTTGAACAGAACAA 370
301 AGCAGCAAGCGATTTTGAACAGAGACGCTTAAAGAAAAGTTTGAACAGAACAA 360
371 GCGATTTGAACAGAGACGCTGCTTAAAGAAAAGTTTGAACAGAACAAAGCGATTAG 430
361 GCGATTTGAACAGAGACGCTGCTTAAAGAAAAGTTTGAACAGAACAAAGCGATTAG 420
431 AACAGAGAGACTGCTTAAAGAAAAGTTTGAACAGAACAAAGCGATTTGAACAGAGA 490
421 AACAGAGAGACTGCTTAAAGAAAAGTTTGAACAGAACAAAGCGATTTGAACAGAGA 480
491 GACGTGCTTAAAGAAAAGTTTGAACAGAACAAAGCGATTTGAACAGAGACGCTGTA 550
481 GACGTGCTTAAAGAAAAGTTTGAACAGAACAAAGCGATTTGAACAGAGACGCTGTA 540
551 AAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGACGCTTAAAGAAAAGT 610
541 AAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGACGCTTAAAGAAAAGT 600
611 TGCAAGACGAGCAAGAGATTTAGAACAAAGAGCGTGTACGAAAAAAATTTAGAAA 670
601 TGCAAGACGAGCAAGAGATTTAGAACAAAGAGCGTGTACGAAAAAAATTTAGAAA 660
671 GAAAAAAGGAAACATGAGATATATTAGCAGAGATTTATATGCTTTAGAAATACAG 730
661 GAAAAAAGGAAACATGAGATATATTAGCAGAGATTTATATGCTTTAGAAATACAG 720
731 CTATGAACTTCATCAGAAAAATGAACGTGATATTTATATACCAATCTTCTTTAC 790
721 CTATGAACTTCATCAGAAAAATGAACGTGATATTTATATACCAATCTTCTTTAC 780
791 CTCAGAGCAACAGAGGAAATAGATAGATTTCCAGAGAAATATCTTAATAGAAAAACA 850
781 CTCAGAGCAACAGAGGAAATAGATAGATTTCCAGAGAAATATCTTAATAGAAAAACA 840
851 ATAGAGAAATCTATTCAACAAATGTTGAAGACGAAAGGATATATCAATAAGACATCTTG 910
841 ATAGAGAAATCTATTCAACAAATGTTGAAGACGAAAGGATATATCAATAAGACATCTTG 900
911 AAAGAAAAGGAAATGTTCAATTAACCAAGAACAAAGAAAGAAATTAATCT 960
901 AAAGAAAAGGAAATGTTCAATTAACCAAGAACAAAGAAAGAAATTAATCT 950

RESULT 8
US-09-900-963-32
Sequence 32, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-900-963-32
Query Match 64.1%; Score 950; DB 3; Length 950;
Best Local Similarity 100.0%; Pred. No. 1,2e-122; Indels 0; Gaps 0;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAACGATCTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATT 70
DB 1 AAACGATCTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATT 60
QY 71 TAGAACAGATAGACTTCTTAAGAAAAGTTTCAAGAGACCAAGCATTTTAACAAG 130
DB 61 TAGAACAGATAGACTTCTTAAGAAAAGTTTCAAGAGACCAAGCATTTTAACAAG 120
QY 131 AGAGACTGTAAAGAAAAGTTGCAAGAACAAACGCGATCTAGAACAAAGAGAGACGTG 190
DB 121 AGAGACTGTAAAGAAAAGTTGCAAGAACAAACGCGATCTAGAACAAAGAGACGTG 180
QY 191 CTAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGAGACGTGCTAAAGAA 250
DB 181 CTAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGAGACGTGCTAAAGAA 240
QY 251 AGTTGCAAGAACAAACGCGATTTAGAACAAAGATAGACTTGTAAAGAAAAGTTACAG 310

DB 241 AGTTGCAAGAACAAACGCGATTTAGAACAAAGATAGACTTGTAAAGAAAAGTTACAG 300
QY 311 AGCAGCAAGCGATTTAGAACAAAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACAA 370
DB 301 AGCAGCAAGCGATTTAGAACAAAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACAA 360
QY 371 GCGATTTAGAACAAAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAG 430
DB 361 GCGATTTAGAACAAAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAG 420
QY 431 AACAGAGAGACTTGTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGA 490
DB 421 AACAGAGAGACTTGTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGA 480
QY 491 GACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGACGTGCTA 550
DB 481 GACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGACGTGCTA 540
QY 551 AAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGACGTGCTAAAGAAAAGT 610
DB 541 AAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGACGTGCTAAAGAAAAGT 600
QY 611 TGCAGAGCAGCAAGAGATTTAGAACAAAGAGCCTGATACGAAAAAATTTAGAAA 670
DB 601 TGCAGAGCAGCAAGAGATTTAGAACAAAGAGCCTGATACGAAAAAATTTAGAAA 660
QY 671 GAAAAAGCAACTGAGATATTTTGCAGAGATTTATATGTCGTTTGAATATCCAG 730
DB 661 GAAAAAGCAACTGAGATATTTTGCAGAGATTTATATGTCGTTTGAATATCCAG 720
QY 731 CTATAGAACTTCATCAGAAAAATGAACGTGATATTATATACCAATCTCTCTTTAC 790
DB 721 CTATAGAACTTCATCAGAAAAATGAACGTGATATTATATACCAATCTCTCTTTAC 780
QY 791 CTCAGAGCAACAGAGGAATAGTAGAGATTCAGAGAAATATCTATATAGAAAAACAA 850
DB 781 CTCAGAGCAACAGAGGAATAGTAGAGATTCAGAGAAATATCTATATAGAAAAACAA 840
QY 851 ATAGAGATTTATTAACAACAATGTTGAGAGACGAAGGATATATCATTAAGGACATCTTG 910
DB 841 ATAGAGATTTATTAACAACAATGTTGAGAGACGAAGGATATATCATTAAGGACATCTTG 900
QY 911 AAGAAAAGAAAGATGGTTCAATTAACCCAGAACAAAAGATTAATCT 960
DB 901 AAGAAAAGAAAGATGGTTCAATTAACCCAGAACAAAAGATTAATCT 950
RESULT 9
US-09-837-344-34
Sequence 34, Application US/09837344
Patent No. US2002004182A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001

```
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/462,625
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: FR 91 01286
/ FILING DATE: 05-FEB-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McGowan, Malcolm K.
/ REGISTRATION NUMBER: 39,300
/ REFERENCE/DOCKET NUMBER: 010830-078
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 34:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 988 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ PUBLICATION INFORMATION:
/ DOCUMENT NUMBER: WO 92/13884
/ PUBLICATION DATE: 20-AUG-1992
/ SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-837-344-34

Query Match      31.9%; Score 472.8; DB 3; Length 988;
Best Local Similarity 95.7%; Pred. No. 1.7e-56;
Matches 486; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 77 AAGATGACTTGTCTAAAGAAAGTTTCAAGAGCAGCAAGCGATTGTAGAACAGAGAGAC 136
DB 481 AAGAAACACGCTAAAGAAAGTTTCAAGAGCAGCAAGCGATTGTAGAACAGAGAGAC 540
QY 137 TTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAGAGAGAGAGAG 196
DB 541 GTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAGAGAGAGAGAG 600
QY 197 AAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAG 256
DB 601 AAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 257 AAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
DB 661 AAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 317 AAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 376
DB 721 AAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 780
QY 377 TAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
DB 781 TAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 437 AGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAG 496
DB 841 AGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAG 900
QY 497 CTAAAGAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAG 556
DB 901 CTAAAGAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAG 960
QY 557 AGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584
DB 961 GGTTCAGAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 988
```

```
/ DRULIHE, Pierre
/ TITLE OF INVENTION: HEPATIC SEQUENCES SPECIFIC FOR THE
/ HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
/ OF STIMULATING THE T LYMPHOCYTES
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
/ STREET: P.O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/900,963
/ FILING DATE: 10-Jul-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/098,327
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McGowan, Malcolm K.
/ REGISTRATION NUMBER: 39,300
/ REFERENCE/DOCKET NUMBER: 010830-045
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 34:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 988 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ PUBLICATION INFORMATION:
/ DOCUMENT NUMBER: WO 92/13884
/ PUBLICATION DATE: 20-AUG-1992
/ SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-900-963-34

Query Match      31.9%; Score 472.8; DB 3; Length 988;
Best Local Similarity 95.7%; Pred. No. 1.7e-56;
Matches 486; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 77 AAGATGACTTGTCTAAAGAAAGTTTCAAGAGCAGCAAGCGATTGTAGAACAGAGAGAC 136
DB 481 AAGAAACACGCTAAAGAAAGTTTCAAGAGCAGCAAGCGATTGTAGAACAGAGAGAC 540
QY 137 TTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGAGAGAGAGAGAGAG 196
DB 541 GTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAG 600
QY 197 AAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 256
DB 601 AAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 257 AAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
DB 661 AAGAACAAAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 317 AAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 376
DB 721 AAGCGATTGTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 780
QY 377 TAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
DB 781 TAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 437 AGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGTAGAACAGAGAGAGAGAG 496
```

Db 841 AGAAGCTTGTAAAGAAAAGTTACAGAGCAGCAAGGATTTTGAACAGATGACCTTG 900
Qy 497 CTAAGAAGAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCGTCTAAAGAA 556
Db 901 CTAAGAAGAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCGTCTAAAGAA 960
Qy 557 AGTTGCAAGAACAAAGCGATTTTGA 584
Db 961 GGTTCAGAAAGAACAAAGCGATTTTGA 988

RESULT 11
US-09-837-344-37
; Sequence 37, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; OF STIMULATING THE T LYMPHOCYTES
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837.344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..954
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-837-344-37

Query Match 31.8%; Score 470.8; DB 3; Length 954;
Best Local Similarity 95.7%; Pred. No. 3.2e-56;
Matches 484; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
77 AAGTAGCTTGTAAAGAAAAGTTACAGAGCAGCAAGGATTTTGAACAGAGAGAC 136

Db 449 AAGAAAAGAGCGTAAAGAAAAGTTACAGGGGCAACAAGCGATTCAGAACAGAGAC 508
Qy 137 TTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTCTAAAG 196
Db 509 GTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTCTAAAG 568
Qy 197 AAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCGTCTAAAGAAAGTTGC 256
Db 569 AAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCGTCTAAAGAAAGTTGC 628
Qy 257 AAGAACAAAGCGATTTTGAACAGAGATGACTTGTCTAAAGAAAAGTTACAGAGCAGC 316
Db 629 AAGAACAAAGCGATTTTGAACAGAGAGATGACTTGTCTAAAGAAAAGTTGCAAGAAC 688
Qy 317 AAAGCGATTTTGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 376
Db 689 AAAGCGATTTTGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 748
Qy 377 TAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAG 436
Db 749 TAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAG 808
Qy 437 AGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCGTG 496
Db 809 AGAGACTTGTCTAAAGAAAAGTTTACAGAGAGAGAGCGATTTTGAACAGATGACTTG 868
Qy 497 CTAAGAAGAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCGTCTAAAGAA 556
Db 869 CTAAGAAGAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCGTCTAAAGAA 928
Qy 557 AGTTGCAAGAACAAAGCGATTTTGA 582
Db 929 GGTTCAGAAAGAACAAAGCGATTTTGA 954

RESULT 12
US-09-900-963-37
; Sequence 37, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; DRUIHE, Pierre
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 37:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..954
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-Aug-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-900-963-37

Query Match      31.8%; Score 470.8; DB 3; Length 954;
Best Local Similarity 95.7%; Pred. No. 3.2e-56;
Matches 484; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 77 AAGATGACTTGCTAAAGAAAGTTACAGAGACGCAAGCGATTAGAAACAAGAGAC 136
DB 449 AAGAAAAACAGCTTAAGAAAAAGTTACAGGGGCAAAAGCGATTACAGAAACAAGAGAC 508
QY 137 TTGCTAAAGAAAAAGTTCAAGAACAAAGCGATCTAGAACAGAGACGCTGCTAAAG 196
DB 509 GTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGACGCTGCTAAAG 568
QY 197 AAAAGTTGCAAGAACAAAGCGATTGAAACAAGAGACGCTGCTAAAGAAAAAGTTGC 256
DB 569 AAAAGTTGCAAGAACAAAGCGATTGAAACAAGAGACGCTGCTAAAGAAAAAGTTGC 628
QY 257 AAGAACCAACAAAGCGATTGAAACAAGATGACCTGCTAAAGAAAAAGTTACAGAGAC 316
DB 629 AAGAACCAACAAAGCGATTGAAACAAGAGACCTGCTAAAGAAAAAGTTACAGAGAC 688
QY 317 AAAGCATTTAGAACAGAGAGACGCTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATT 376
DB 689 AAAGCATTTAGAACAGAGAGACGCTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATT 748
QY 377 TAGAACCAAGAGACGCTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATTGAAACAAG 436
DB 749 TAGAACCAAGAGACGCTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATTGAAACAAG 808
QY 437 AAGACCTTGCTAAAGAAAAAGTTGCAAGAACCAAGAGACGATTAGAACAGAGACGCTG 496
DB 809 AAGACCTTGCTAAAGAAAAAGTTGCAAGAGCAGCAAGCGATTGAAACAAGATGACCTG 868
QY 497 CTAAGAAAAAGTTGCAAGAACCAAGCGATTAGAACAGAGACGCTGCTAAAGAAAA 556
DB 869 CTAAGAAAAAGTTGCAAGAACCAAGCGATTAGAACAGAGACGCTGCTAAAGAAAA 928
QY 557 AGTTGCAAGAACCAAGCGATTTA 582
DB 929 GGTTCAGAGAACCAAGCGATTTA 954

RESULT 13
US-10-668-749A-1
; Sequence 1, Application US/10668749A
; Publication No. US20040110205A1
; GENERAL INFORMATION:
; APPLICANT: Agilent Technologies
; TITLE OF INVENTION: Methods and Systems for Nanopore Data Analysis
; FILE REFERENCE: 50112-1580
; CURRENT APPLICATION NUMBER: US/10/668,749A
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
OTHER INFORMATION: synthetic construct
US-10-668-749A-1

Query Match      15.7%; Score 233.4; DB 7; Length 1300;
Best Local Similarity 48.7%; Pred. No. 2.8e-23;
Matches 633; Conservative 0; Mismatches 666; Indels 0; Gaps 0;

QY 91 AAAGAAAAAGTTACAGAGACGCAAAAGCGATTAGAACAGAGAGACTTGTAAAGAAAAAG 150
DB 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60
QY 151 TTGCAAGAACCAAGCGATCTAGAACAGAGACGCTGCTAAAGAAAAAGTTGCAAGAAC 210
DB 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120
QY 211 CAACAAACGATTAGAACAGAGAGAGCTGCTAAAGAAAAAGTTGCAAGAACCAAGAAC 270
DB 121 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180
QY 271 GATTAGAACCAAGATGACCTGCTAAAGAAAAAGTTACAGAGACGCAAAAGCGATTAGAAC 330
DB 181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 240
QY 331 CAAGAGACGCTGCTAAAGAAAAAGTTGCAAGAACCAAGCGATTGAAACAAGAGAC 390
DB 241 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300
QY 391 CGTGTAAAGAAAAAGTTCAAGAACCAAGCGATTAGAACAGAGACGCTGCTAAAG 450
DB 301 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 360
QY 451 GAAAGATTGCAAGAACCAAGCGATTAGAACAGAGACGCTGCTAAAGAAAAAGTTG 510
DB 361 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420
QY 511 CAAGAACCAAGCGATTGAAACAAGAGAGACGCTGCTAAAGAAAAAGTTGCAAGAAC 570
DB 421 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 480
QY 571 CAAGCGATTAGAACAGAGAGACGCTGCTAAAGAAAAAGTTGCAAGAGCAGCAAGAGAT 630
DB 481 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 540
QY 631 TTAGAACCAAGAGAGCTGATCGAATAAATTTAGAAAGAAAAAGAAACATGAGAT 690
DB 541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600
QY 691 ATATTAGAGAGATTATATGCTGTTAGAAATACAGCTATAGAACTTCATCAGAA 750
DB 601 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 660
QY 751 AATGAACGTGATATTATATACCATCAATCTTTTACCTCAGAGACAGAGGAGAT 810
DB 661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 720
QY 811 AGTAGAGATTCCAGGAATATCTATTAATAGAAAAACAATATAGAAATCTATTACAACA 870
DB 721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 780
QY 871 AATGTTAAGAGCAGAGGATATACATAGAGACATCTTGAAAGAAAGAGAGATGCTTCA 930
DB 781 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 840
QY 931 AATAAACCAAGAACCAAGAAAGATATCTGCTGATACATCAAAATCATATTAGAGACA 990
DB 841 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900
QY 991 GTAATATTCTGATGTTAATGATTTTCAATTAAGATATGAGATGAGATGATGCT 1050
DB 901 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 960
QY 1051 GAATATGACGATTCTTAATAGATGAGAGAGATGATGAAGACTTAGCAATTTAAG 1110
```

Db 961 AA 1020
Qy 1111 CTTATGTGCAATATGCAATTTCCAGATGAGAGAAACATAGAAATTTATAAAGACTA 1170
Db 1021 AA 1080
Qy 1171 GAAGATTGATAGAGAAATGAAATTTAGATGATTTAGATGAGAAATAGAAATCA 1230
Db 1081 AA 1140
Qy 1231 TCAGAGAAATTTATCTGAGAGAAATTAATAAGAGAAATATGAGAAATCAAGAT 1290
Db 1141 AA 1200
Qy 1291 AATTAATTTAAACCAATATGATTAAGTTGTATGATGACATTTAAATATTAAT 1350
Db 1201 AA 1260
Qy 1351 GATAGAGAGTTATATAGAGAAAGAGAAATTCATAAA 1389
Db 1261 AA 1299

RESULT 14

US-10-425-115-67687/c
; Sequence 67687, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Le Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 67687
; LENGTH: 1297
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1297)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: clone ID: MRT4577_161728C.1
US-10-425-115-67687

Query Match 10.3%; Score 153.2; DB 8; Length 1297;
Best Local Similarity 47.2%; Pred. No. 3.6e-12;
Matches 481; Conservative 0; Mismatches 538; Indels 1; Gaps 1;

Qy 5 AACACAAAGCGATCTGACACAGAGACGCTGCTAAAGAAAGTTGCAAGACAAACAA 64
Db 1259 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
Qy 65 GCGATTGAGACAGATGAGCTTCTAAAGAAAGTTACAGAGACAGCAAGGATTTAG 124
Db 1199 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
Qy 125 AACAGAGAGACTTGTAAAGAAAGTTGCAAGACAAACAGGATCTAGAACAGAGA 184
Db 1139 AAAAAACGAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
Qy 185 GAGCTGCTAAAGAAAGTTGCAAGACAAACAGGATTTGAAACAGAGAGCTGCTA 244
Db 1079 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
Qy 245 AAGAAAGTTGCAAGACAAACAGGATTTAGAACAGATGAGCTTGTAAAGAAAGT 304
Db 1019 AAAAAAAGAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 960

Qy 305 TACAGAGCAGCAAGCGATTTAGAACAGAGACGCTGTAAAGAAAGTTGCAAGAAC 364
Db 959 AA 900
Qy 365 AACAGAGCGATTTGAGACAGAGAGCGTCTTAAGAAAGTTGCAAGACAAACAGCG 424
Db 899 AA 840
Qy 425 ATTTAGAACAGAGAGACTTGTAAAGAAAGTTGCAAGACAAACAGGATTTAGAAC 484
Db 839 AA 780
Qy 485 AAGAGAGCGCTGCTAAAGAAAGTTGCAAGACAAACAGGATTTAGAACAGAGAGA 543
Db 779 AA 720
Qy 544 CGTGTAAAGAAAGTTGCAAGACAAACAGGATTTAGAACAGAGAGCGTGTAAAC 603
Db 719 AAAAAACAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
Qy 604 GAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGGCTGATACGAAAAAAT 663
Db 659 AAAAAAAGAAAAAAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
Qy 664 TTGAGAGAAAAAAGAGACATGAGATATTTGACAGAGATTTATATGCTGTTTGA 723
Db 599 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
Qy 724 ATACAGCTATAGACTTCATCAGAAATGAACGTGATATTAATTAATTAATTAATTA 783
Db 539 AA 480
Qy 784 TCTTTAAGTACAGACAGAGGAAATAGTAGATTTCCAGAGAAATATCTATATAGAA 843
Db 479 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
Qy 844 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 903
Db 419 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
Qy 904 CATCTGAGAGAAAGAGAGGTTGCTCAATTAATTAATTAATTAATTAATTAATTC 963
Db 359 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
Qy 964 GACATCAAAATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1023
Db 299 CACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240

RESULT 15

US-10-221-714A-461/c
; Sequence 461, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: Olek, Alexander
; APPLICANT: PIEPENBROCK, Christien
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 540
SEQ ID NO 461
LENGTH: 6292
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-461

Query Match 10.2%; Score 151.2; DB 7; Length 6292;
Best Local Similarity 46.2%; Pred. No. 9.2e-12;
Matches 619; Conservative 0; Mismatches 708; Indels 13; Gaps 3;

QY 40 AAAAAAAAAAGTTCAGACAGACACAAACCGATTAGAACAGATGACTTCTTAAGAAAAG 99
DB 4548 AA 4489
QY 100 TTACAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 159
DB 4488 CTTAAAAACGATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4429
QY 160 CACAAAGCGATCTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAGAACACAAAGC 219
DB 4428 AA 4369
QY 220 GATTAGAACAGAGAGACGCTGCTAAAGAAAAGTTGCAGAACACAAAGCATTAGAA 279
DB 4368 TTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4309
QY 280 CAAGATAGACTGCTTAAGAAAAGTTACAAGACGACGACGACGACGACGACGACGACGAC 335
DB 4308 AAAAAAAAAATTTACAAAATATATATATATATATATATATATATATATATATATAT 4249
QY 336 GAGACGCTGCTAAAGAAAAGTTGCAGAACACAAACGATTAGAACCAAGAGACGCTGC 395
DB 4248 AAAAAAAAAATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4189
QY 396 TAAAGAAAAGTTGCAGAACACAAAGCATTAGAACAGAGACGCTGCTTAAGAAAAG 455
DB 4188 CAAA 4129
QY 456 GTTGAAGAACACAAAGCATTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAGAAC 515
DB 4128 CATTAAAAAATACTTTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4069
QY 516 ACACAAAGCGATTTAGAAC---AAGAGACGCTGCTAAAGAAAAGTTGCAGAACAC 571
DB 4068 AATAAAAAAAAAATTAAGAACCGAAAAAATAAATTTTAACTAATAAAGATTAACAAA 4009
QY 572 AAAAGCATTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAGAGCAGCAAGAGATT 631
DB 4008 AAAAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3949
QY 632 TAGAACAAAGAGCGCTGATAGCAAAAAAATTTAGAAAGAAAAAGAACATGAGATA 691
DB 3948 AAAAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAAC 3889
QY 692 TATTAGCAGAGATTATATGCTGTTTGAATAACAGCTATAGAACTTCATCAGAAA 751
DB 3888 TAAAAAAAAAAGATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 3829
QY 752 ATGAACGCTGATATTATATCCACATCATCTTCTTAAGCAAGCAAGAGGGAATA 811
DB 3828 AACAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3769
QY 812 GTAGAGATTCAGAGAAATATCTATAATAGAAAAAATAAGAAATCTATTACAA 871
DB 3768 CAAAAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3709
QY 872 ATGTTGAGAGACGAGAGATATATCATTAAGACATCTTGAAGAAAAGAAAGATGTTCAA 931

DB 3708 AACTTACAAAAAAGCAAAAAAATATTTTTTAAATAATAAAAAAAAAAAAAATAA 3649
QY 932 TAAACGAGACAAAAAGATTAATCTGCTGACATACAAAATCATACATTAGAGACAG 991
DB 3648 CAAAAAAAAAAAAAAAAAAAAAAGATTAATAATAATAATAATAATAATAATAATAATA 3589
QY 992 TAAATATTTCTGATGTTATGATTTTCAATTAAGTATGAGATGAATAAGTGCT- 1050
DB 3588 AAAAAAAAAACAAACAAAAATTAATAATAATAATAATAATAATAATAATAATAATA 3529
QY 1051 ----GAATATGACATTCATTATATGATGAAGAAAGATGATGAAGCTTGAAGATT 1106
DB 3528 TTAATAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3469
QY 1107 TAGCCTATTGTGCAATATGACAAATTTCCAGATGAAGAAACATAGCAATTTTAAGA 1166
DB 3468 AAAAAAAAAAAAAAACAAGATTAATAATAATAATAATAATAATAATAATAATAA 3409
QY 1167 ACTAGAACATTTGATAGAAAAATGAATAATTTAGATGATTTAGATGAAGATAGAAA 1226
DB 3408 AATTAACGATTAATATATTAATAAATACTTACAAAAAACAATTAATAATAATAA 3349
QY 1227 ATCATGAGAAATTTCTGAGAAAAATTAATAATAATAATAATAATAATAATAATA 1286
DB 3348 AAAAAATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 3289
QY 1287 GATATATATTTTAAACCAATGATTAATAAGTTGTATGATGAGCATATTAAATAATA 1346
DB 3288 GAAAAAATAATTTCTCTCTTACTATTTTAACTCTTCAATTCGATCAAAAAA 3229
QY 1347 AATGATTAAGCAGTTAATA 1366
DB 3228 TAAATTAATATATTTCAAAA 3209

Search completed: December 7, 2005, 06:33:43
Job time : 1337 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 23:17:43 ; Search time 296 Seconds

(without alignments)
8899.819 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1482

Sequence: 1 CAGAACACACAAAGCGATCT.....AATTTTATGAACTAATA 1482

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1.COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5.COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H.COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCUS.COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1482	100.0	1482	US-08-098-327E-41	Sequence 41, Appl
2	1482	100.0	1482	US-08-462-625-41	Sequence 41, Appl
3	1480.4	99.9	1482	US-08-098-327E-45	Sequence 45, Appl
4	1480.4	99.9	1482	US-08-462-625-45	Sequence 45, Appl
5	1446	97.6	1493	US-08-098-327E-38	Sequence 38, Appl
6	1446	97.6	1493	US-08-462-625-38	Sequence 38, Appl
7	950	64.1	950	US-08-098-327E-32	Sequence 32, Appl
8	950	64.1	950	US-08-462-625-32	Sequence 32, Appl
9	853.2	57.6	1320	US-08-257-073-15	Sequence 15, Appl
10	472.8	31.9	988	US-08-098-327E-34	Sequence 34, Appl
11	472.8	31.9	988	US-08-462-625-34	Sequence 34, Appl
12	470.8	31.8	954	US-08-098-327E-37	Sequence 37, Appl
13	470.8	31.8	954	US-08-462-625-37	Sequence 37, Appl
14	129.8	8.8	1039	US-09-902-540-1280	Sequence 1280, Ap
15	120.8	8.2	612	US-09-902-540-1357	Sequence 1357, Ap
16	116.2	7.8	3095	5231168-1	Patent No. 5231168
17	111.6	7.5	1827	US-09-270-767-1308	Sequence 1308, Ap
18	111.6	7.5	1827	US-09-270-767-16590	Sequence 16590, A
19	110.2	7.4	7218	US-08-232-463-14	Sequence 14, Appl
20	101.2	6.8	1956	US-08-559-896B-1	Sequence 1, Appl
21	101.2	6.8	1956	US-09-351-794A-1	Sequence 1, Appl
22	95	6.4	187169	US-09-949-016-12775	Sequence 12775, A
23	95	6.4	191569	US-09-949-016-15940	Sequence 15940, A
24	94.8	6.4	614	US-09-902-540-1318	Sequence 1318, Ap

25	93.6	6.3	5361	3	US-08-973-462-2	Sequence 2, Appl
26	93.6	6.3	6152	3	US-08-973-462-1	Sequence 1, Appl
27	91.6	6.2	29927	3	US-09-949-016-11814	Sequence 11814, A
28	91.6	6.2	29927	3	US-09-949-016-17474	Sequence 17474, A
29	91.6	6.2	29927	3	US-09-949-016-17475	Sequence 17475, A
30	91.6	6.2	37802	3	US-09-949-016-12639	Sequence 12639, A
31	89.6	6.0	1664976	3	US-08-916-421B-1	Sequence 1, Appl
32	89.6	6.0	1664976	3	US-09-692-570-1	Sequence 1, Appl
33	88.4	6.0	5340	3	US-09-627-122-21	Sequence 21, Appl
34	87	5.9	639	3	US-09-248-796A-9722	Sequence 9722, Ap
35	86.8	5.9	4766	6	PCT-US93-07261-10	Sequence 10, Appl
36	86.4	5.8	601	3	US-09-949-016-203127	Sequence 203127,
37	86.4	5.8	601	3	US-09-949-016-203169	Sequence 203169,
38	85	5.7	147382	3	US-09-949-016-14624	Sequence 14624, A
39	83.2	5.6	16659	3	US-09-461-697-184	Sequence 184, Appl
40	82.8	5.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
41	81.4	5.5	1685	3	US-10-104-047-1957	Sequence 1957, Ap
42	81.4	5.5	2223	2	US-08-257-073-4	Sequence 4, Appl
43	81.2	5.5	696	3	US-09-461-697-193	Sequence 193, App
44	81.2	5.5	699	3	US-09-461-697-191	Sequence 191, App
45	81.2	5.5	717	3	US-09-461-697-189	Sequence 189, App

ALIGNMENTS

RESULT 1
US-08-098-327E-41
; Sequence 41, Application US/08098327E
; Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/33884
PUBLICATION DATE: 20-AUG-1992

US-08-098-337E-41

Query Match 100.0%; Score 1482; DB 3; Length 1482;
Best Local Similarity 100.0%; Pred. No. 1.8e-246;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 CAAAGAACAAAGCGATCTAGAACAGAGACGTCTAAAGAAAGTTGCAAGACAA
DB      1 CAAAGAACAAAGCGATCTAGAACAGAGACGTCTAAAGAAAGTTGCAAGACAA
QY      61 CAAACGATTTGAAACAAGATAGACTTGTAAAGAAAGTTACAGAGACAGAAAGCAT
DB      61 CAAACGATTTGAAACAAGATAGACTTGTAAAGAAAGTTACAGAGACAGAAAGCAT
QY      121 TTAGAACAAAGAGAGACTTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAA
DB      121 TTAGAACAAAGAGAGACTTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAA
QY      181 GAGAGACGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGT
DB      181 GAGAGACGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGT
QY      241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGCTAAAGAA
DB      241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGCTAAAGAA
QY      301 AAGTTACAAAGACAGACGATTTAGAACAAAGAGAGACGTCTAAAGAAAGTTGCA
DB      301 AAGTTACAAAGACAGACGATTTAGAACAAAGAGAGACGTCTAAAGAAAGTTGCA
QY      360 AAGTTACAAAGACAGACGATTTAGAACAAAGAGAGACGTCTAAAGAAAGTTGCA
DB      360 AAGTTACAAAGACAGACGATTTAGAACAAAGAGAGACGTCTAAAGAAAGTTGCA
QY      420 GAAACAACAAAGCGATTTAGAACAAAGAGAGACGTCTAAAGAAAGTTGCAAGACAA
DB      420 GAAACAACAAAGCGATTTAGAACAAAGAGAGACGTCTAAAGAAAGTTGCAAGACAA
QY      480 ACGGATTTGAAACAAAGAGACTTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTA
DB      480 ACGGATTTGAAACAAAGAGACTTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTA
QY      540 GAAACAAGAGAGACGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG
DB      540 GAAACAAGAGAGACGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG
QY      600 AGACGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGTGCT
DB      600 AGACGTCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGTGCT
QY      660 AAAAGAAAGTTGCAAGAGCAAGCAAGAAAGTTTGAACAAAGAGAGCGTATCGAAAAA
DB      660 AAAAGAAAGTTGCAAGAGCAAGCAAGAAAGTTTGAACAAAGAGAGCGTATCGAAAAA
QY      720 AATTGAAAGAAAGAAAGAAAGCAATGAGATATTTAGAGAGAGATTTATATGTCGTTA
DB      720 AATTGAAAGAAAGAAAGAAAGCAATGAGATATTTAGAGAGAGATTTATATGTCGTTA
QY      780 GAAATACCAAGCTATAGAACTTCATCAGAAAAATGAACGTGATATTAATACCACTCAA
DB      780 GAAATACCAAGCTATAGAACTTCATCAGAAAAATGAACGTGATATTAATACCACTCAA
QY      840 TCTTTCTTTTACCTCAGACAAACAGAGGAATAGTAGAGATTTCCAGAGAAATATCTATAA
DB      840 TCTTTCTTTTACCTCAGACAAACAGAGGAATAGTAGAGATTTCCAGAGAAATATCTATAA
QY      900 GAAAAAACAATAGAGATCTATTACAACAAATGTTAGAGAGAGAGGATTAACATAA
DB      900 GAAAAAACAATAGAGATCTATTACAACAAATGTTAGAGAGAGAGGATTAACATAA
QY      960 GGCATCTTTGAGAAAAAGAAAGATGTTCAATTAACAAAGAAACAAAGAAAGATTAATCT
DB      960 GGCATCTTTGAGAAAAAGAAAGATGTTCAATTAACAAAGAAACAAAGAAAGATTAATCT

```

```

QY      961 GCTGACATCAAAATCATATACATTAGAGACAGTAATAATTTCTGATGTTAAGATTTCAA
DB      961 GCTGACATCAAAATCATATACATTAGAGACAGTAATAATTTCTGATGTTAAGATTTCAA
QY      1021 ATTAAGTAAGTATGAGATGAAATATAGTCTGAATATGACATTCATTATATGATGAAGAA
DB      1021 ATTAAGTAAGTATGAGATGAAATATAGTCTGAATATGACATTCATTATATGATGAAGAA
QY      1080 GAAATATGAGAGAGCTTAAGAGAAATTTAAGCTTATGCTAATATGACAAATTTCCAGAT
DB      1080 GAAATATGAGAGAGCTTAAGAGAAATTTAAGCTTATGCTAATATGACAAATTTCCAGAT
QY      1140 GAAATATGAGAGAGCTTAAGAGAAATTTAAGAGATTTGATAGAGAAATGAAATTTA
DB      1140 GAAATATGAGAGAGCTTAAGAGAAATTTAAGAGATTTGATAGAGAAATGAAATTTA
QY      1200 GATGATTTAGATGAGAGATGAAATATCATCAGAGAAATTTATCTGAAGAAAAATATAA
DB      1200 GATGATTTAGATGAGAGATGAAATATCATCAGAGAAATTTATCTGAAGAAAAATATAA
QY      1260 GATGATTTAGATGAGAGATGAAATATCATCAGAGAAATTTATCTGAAGAAAAATATAA
DB      1260 GATGATTTAGATGAGAGATGAAATATCATCAGAGAAATTTATCTGAAGAAAAATATAA
QY      1320 AAAGAAAGAAATATGAGAGAAATTTAAGAGATTTAATTAACCAATGATTAAGATTTG
DB      1320 AAAGAAAGAAATATGAGAGAAATTTAAGAGATTTAATTAACCAATGATTAAGATTTG
QY      1380 AAAGAAAGAAATATGAGAGAAATTTAAGAGATTTAATTAACCAATGATTAAGATTTG
DB      1380 AAAGAAAGAAATATGAGAGAAATTTAAGAGATTTAATTAACCAATGATTAAGATTTG
QY      1440 TTGATGAGCATATTTAAAGAAATATTAAGATTAAGAGATTTAAGAGATTTAAGAGATTTA
DB      1440 TTGATGAGCATATTTAAAGAAATATTAAGATTAAGAGATTTAAGAGATTTAAGAGATTTA
QY      1482 GAGTTATCTGAAGATATTAATATTTTGAAGATTTA
DB      1482 GAGTTATCTGAAGATATTAATATTTTGAAGATTTA

```

RESULT 2

US-08-462-625-41

/ Sequence 41, Application US/08462625

/ Patent No. 6319502

/ GENERAL INFORMATION:

/ APPLICANT: GUERIN-MARCHAND, Claudine

/ APPLICANT: DRULIHE, Pierre

/ TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

/ TITLE OF INVENTION: HEPATIC STAGES OF P. PALCIPARUM BEARING EPITOPES CAPABLE

/ OF STIMULATING THE T LYMPHOCYTES

/ NUMBER OF SEQUENCES: 46

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Burns, Doane, Swecker & Mathis

/ STREET: P. O. Box 1404

/ CITY: Alexandria

/ STATE: Virginia

/ COUNTRY: United States

/ ZIP: 2213-1404

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/462,625

/ FILING DATE: 05-JUN-1995

/ CLASSIFICATION: 424

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 08/098,327

/ FILING DATE: 24-NOV-1993

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: FR 91 01286

/ FILING DATE: 05-FEB-1991

/ ATTORNEY/AGENT INFORMATION:

/ NAME: McGowan, Malcolm K.

/ REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-41

Query Match 100.0%; Score 1482; DB 3; Length 1482;
Best Local Similarity 100.0%; Pred. No. 1,8e-246;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGAACAAACGATCTTGAACAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAA 60
DB 1 CAAGAACAAACGATCTTGAACAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAA 60
QY 61 CAAAGCATTTAGAACAGATGACTTCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGAT 120
DB 61 CAAAGCATTTAGAACAGATGACTTCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGAT 120
QY 121 TTAGAACAGAGAGACTTGTCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATCTTGAACAA 180
DB 121 TTAGAACAGAGAGACTTGTCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATCTTGAACAA 180
QY 181 GAGAGAGCTGCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATCTTGAACAA 240
DB 181 GAGAGAGCTGCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATCTTGAACAA 240
QY 241 GCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATCTTGAACAAAGCAAGCAAGCAAG 300
DB 241 GCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATCTTGAACAAAGCAAGCAAGCAAG 300
QY 301 AAGTTACAGAGAGCAAGCAAGCGATTTAGAACAGAGAGAGCGTCTTAAAGAAAAGTTGCA 360
DB 301 AAGTTACAGAGAGCAAGCAAGCGATTTAGAACAGAGAGAGCGTCTTAAAGAAAAGTTGCA 360
QY 361 GAAACAAACAAAGCGATTTAGAACAGAGAGAGCGTCTTAAAGAAAAGTTGCAAGCAAGCA 420
DB 361 GAAACAAACAAAGCGATTTAGAACAGAGAGAGCGTCTTAAAGAAAAGTTGCAAGCAAGCA 420
QY 421 AGCGATTTAGAACAGAGAGAGCGTCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATTTA 480
DB 421 AGCGATTTAGAACAGAGAGAGCGTCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATTTA 480
QY 481 GAACAGAGAGAGCGTCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATTTAGAACAGAG 540
DB 481 GAACAGAGAGAGCGTCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATTTAGAACAGAG 540
QY 541 AGAGCTGCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATTTAGAACAGAGAGCGTCT 600
DB 541 AGAGCTGCTTAAAGAAAAGTTGCAAGAGCAAGCAAGCGATTTAGAACAGAGAGCGTCT 600
QY 601 AAAGAAAAGTTGCAAGAGCAAGCAAGCGATTTAGAACAGAGAGCGTCTTAGCAAGAAA 660
DB 601 AAAGAAAAGTTGCAAGAGCAAGCAAGCGATTTAGAACAGAGAGCGTCTTAGCAAGAAA 660
QY 661 AATTAGAAAAGAAAAGCAATGAGATATATTAGCAGAGAGATTTATATGTCGTTTA 720
DB 661 AATTAGAAAAGAAAAGCAATGAGATATATTAGCAGAGAGATTTATATGTCGTTTA 720
QY 721 GAATATCAAGCTATGAACTTCCATCAAGAAAATGAACCTGATATTTATATCCACATCAA 780
DB 721 GAATATCAAGCTATGAACTTCCATCAAGAAAATGAACCTGATATTTATATCCACATCAA 780

DB 721 GAATATCAAGCTATGAACTTCCATCAAGAAAATGAACCTGATATTTATATCCACATCAA 780
QY 781 TCTTCTTTACTAGAGCAACAGAGGAAATGATGAGATTTCCAGAGAAAATCTATATATA 840
DB 781 TCTTCTTTACTAGAGCAACAGAGGAAATGATGAGATTTCCAGAGAAAATCTATATATA 840
QY 841 GAAAAAACAATAGAGATCTATTTACAAACAATGTTGAAGAGCAAGAGATATATCATATA 900
DB 841 GAAAAAACAATAGAGATCTATTTACAAACAATGTTGAAGAGCAAGAGATATATCATATA 900
QY 901 GGAATCTTGAAGAAAAGAAAGTGTCTCATATAAACAGAACAAAAAGAGATTAATCT 960
DB 901 GGAATCTTGAAGAAAAGAAAGTGTCTCATATAAACAGAACAAAAAGAGATTAATCT 960
QY 961 GCTGACATTAACAATCATATTTAGAGACAGTAAATTTTCTGATGTTATGATTTTCA 1020
DB 961 GCTGACATTAACAATCATATTTAGAGACAGTAAATTTTCTGATGTTATGATTTTCA 1020
QY 1021 ATTAAGTATGATGAGATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 ATTAAGTATGATGAGATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 GAAGATGATGAGATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 GAAGATGATGAGATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GAAAGAAAATGAGATTTTAAAGAACTTGAAGATTTGATGAGAAAAATGAAAAATTTA 1200
DB 1141 GAAAGAAAATGAGATTTTAAAGAACTTGAAGATTTGATGAGAAAAATGAAAAATTTA 1200
QY 1201 GATGATTTAGATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 GATGATTTAGATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 AAAGAAAAGAAATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 AAAGAAAAGAAATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 TATGATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 TATGATGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 TTTCTTAAATCATTTGTTTCAATATTTGACGAGACATATAATTTTACAGATCTGAT 1440
DB 1381 TTTCTTAAATCATTTGTTTCAATATTTGACGAGACATATAATTTTACAGATCTGAT 1440
QY 1441 GAGTATCTGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1482
DB 1441 GAGTATCTGAGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1482

RESULT 3
US-08-327E-45
Sequence 45, Application US/08098327E
Patent No. 5270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-45

Query Match 99.9%; Score 1480.4; DB 3; Length 1482;
Best Local Similarity 99.9%; Pred. No. 3.5e-246;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAAGAACAAAGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAGACAA 60
DB 1 CAAGAACAAAGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAAGACAA 60
OY 61 CAAGCGATTTAGAACAAAGTGAAGCTTGAAGAAAAGTTGCAAGAGCAAGCGAT 120
DB 61 CAAGCGATTTAGAACAAAGTGAAGCTTGAAGAAAAGTTGCAAGAGCAAGCGAT 120
OY 121 TTAGAACAGAGAGACTTCTAAAGAAAAGTTGCAAGACAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAGAGAGACTTCTAAAGAAAAGTTGCAAGACAAAGCGATCTAGAACAA 180
OY 181 GAGAGACGCTTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAGACGCT 240
DB 181 GAGAGACGCTTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAGACGCT 240
OY 241 GCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGATGACTTGTCTAAAGAA 300
DB 241 GCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGATGACTTGTCTAAAGAA 300
OY 301 AAGTTCAAGAGACGACAAAGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTGCAA 360
DB 301 AAGTTCAAGAGACGACAAAGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTGCAA 360
OY 361 GAAACAACAAGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTGCAAGACAA 420
DB 361 GAAACAACAAGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTGCAAGACAA 420
OY 421 AGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTGCAAGACAAAGCGATTTA 480
DB 421 AGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTGCAAGACAAAGCGATTTA 480
OY 481 GAAACAAGAGACGCTCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAG 540
DB 481 GAAACAAGAGACGCTCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAG 540
OY 541 AGACGCTGCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAGACGCTGCT 600
DB 541 AGACGCTGCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAGACGCTGCT 600

OY 601 AAAGAAAAGTTGCAAGACAGACAAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAA 660
DB 601 AAAGAAAAGTTGCAAGACAGACAAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAA 660
OY 661 AATTTAGAAAAGAAAAGAAACATGAGATATATTAGCAGAGATTTATATGTCGTTTA 720
DB 661 AATTTAGAAAAGAAAAGAAACATGAGATATATTAGCAGAGATTTATATGTCGTTTA 720
OY 721 GAATACAGCTATATGAACTTCCATCAAGAAATGAGTGAATATTATATACCATCA 780
DB 721 GAATACAGCTATATGAACTTCCATCAAGAAATGAGTGAATATTATATACCATCA 780
OY 781 TCTTCTTACCTCAGACCAAGAGAGATAGATAGATTTCCAGAAATATCTATATA 840
DB 781 TCTTCTTACCTCAGACCAAGAGAGATAGATAGATTTCCAGAAATATCTATATA 840
OY 841 GAAAAACAAATATAGAAATCTATTTCAACAAATGTTGAAAGACGAGATATACATAA 900
DB 841 GAAAAACAAATATAGAAATCTATTTCAACAAATGTTGAAAGACGAGATATACATAA 900
OY 901 GAGACATCTGAAAGAAAAGATGCTTCAATTAACAGAACAAAGAAAGATTAATCT 960
DB 901 GAGACATCTGAAAGAAAAGATGCTTCAATTAACAGAACAAAGAAAGATTAATCT 960
OY 961 GCTGACATACAAAATCATACATTTAGAGACAGTAAATATTTCTGATTTAATGATTTCA 1020
DB 961 GCTGACATACAAAATCATACATTTAGAGACAGTAAATATTTCTGATTTAATGATTTCA 1020
OY 1021 ATAGATTAATGATGAGATGAAATTAAGTCTGCAATATGACATTTCAATATAGATAAGAA 1080
DB 1021 ATAGATTAATGATGAGATGAAATTAAGTCTGCAATATGACATTTCAATATAGATAAGAA 1080
OY 1081 GAAAGTGAATGAAAGCTTGAAGAAATTTAAGCTATGTCATATAGACATTTCCAGAT 1140
DB 1081 GAAAGTGAATGAAAGCTTGAAGAAATTTAAGCTATGTCATATAGACATTTCCAGAT 1140
OY 1141 GAAGAAAACATAGAAATTTATAGAACTAGAGATTTGATAGAGAAAAATGAATTTTA 1200
DB 1141 GAAGAAAACATAGAAATTTATAGAACTAGAGATTTGATAGAGAAAAATGAATTTTA 1200
OY 1201 GATGATTTAGATGAAGAAATGAAAAATCATCAGAGAAATTTATCGAAGAAAAATTA 1260
DB 1201 GATGATTTAGATGAAGAAATGAAAAATCATCAGAGAAATTTATCGAAGAAAAATTA 1260
OY 1261 AAAGAAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTTG 1320
DB 1261 AAAGAAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTTG 1320
OY 1321 TATGATGACATATTAATAAATATTAATAATGTAAGCAGTTAATAGAAAAAGAAAA 1380
DB 1321 TATGATGACATATTAATAAATATTAATAATGTAAGCAGTTAATAGAAAAAGAAAA 1380
OY 1381 TTCAATTAATATCTGTTTCAATATATTGACGAGACAAATGAATTTTACAGATCTGAT 1440
DB 1381 TTCAATTAATATCTGTTTCAATATATTGACGAGACAAATGAATTTTACAGATCTGAT 1440
OY 1441 GAGTATCTGAAGATATTAATTAATTTTGAAGACTATA 1482
DB 1441 GAGTATCTGAAGATATTAATTAATTTTGAAGACTATA 1482

RESULT 4
US-08-462-625-45
Sequence 45, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPS CAPABLE
NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:
ADDRESSEE: Burne, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-45

Query Match 99.9%; Score 1480.4; DB 3; Length 1482;
Best Local Similarity 99.9%; Pred. No. 3.5e-246;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAGAACAAAGCGATCTGAAACAGAGAGCGTCTAAAGAAAAGTTCAGAACAA 60
DB 1 CAAGAACAAAGCGATCTGAAACAGAGAGCGTCTAAAGAAAAGTTCAGAACAA 60
QY 61 CAAAGCGATTAGAACAAAGATGAGCTTGCTAAAGAAAAGTTCAGAACAAAGCGAT 120
DB 61 CAAAGCGATTAGAACAAAGATGAGCTTGCTAAAGAAAAGTTCAGAACAAAGCGAT 120
QY 121 TTAGAACAGAGAGAGCTTCTAAAGAAAAGTTCAGAACAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAGAGAGAGCTTCTAAAGAAAAGTTCAGAACAAAGCGATCTAGAACAA 180
QY 181 GAGAGACGTCTAAAGAAAAGTTCAGAACAAAGCGATTTTGAACAGAGAGAGCT 240
DB 181 GAGAGACGTCTAAAGAAAAGTTCAGAACAAAGCGATTTTGAACAGAGAGAGCT 240
QY 241 GCTAAGAAAAGTTCAGAACAAAGCGATTTAGAACAAAGATGAGCTTGCTAAAGAA 300
DB 241 GCTAAGAAAAGTTCAGAACAAAGCGATTTAGAACAAAGATGAGCTTGCTAAAGAA 300
QY 301 AAGTTACAGAGAGCAAGCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCA 360
DB 301 AAGTTACAGAGAGCAAGCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCA 360
QY 361 GAACAAACAAAGCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 420

DB 361 GAACAAACAAAGCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 420
QY 421 AGCGATTTAGAACAGAGAGCTTGCTAAAGAAAAGTTCAGAACAAAGCGATTTA 480
DB 421 AGCGATTTAGAACAGAGAGCTTGCTAAAGAAAAGTTCAGAACAAAGCGATTTA 480
QY 481 GAACAAAGAGAGCGTCTAAAGAAAAGTTCAGAACAAAGCGATTTAGAACAGAG 540
DB 481 GAACAAAGAGAGCGTCTAAAGAAAAGTTCAGAACAAAGCGATTTAGAACAGAG 540
QY 541 AGAGTGTCTAAAGAAAAGTTCAGAACAAAGCGATTTAGAACAGAGAGAGCGTCT 600
DB 541 AGAGTGTCTAAAGAAAAGTTCAGAACAAAGCGATTTAGAACAGAGAGAGCGTCT 600
QY 601 AAAGAAAAGTTCAGAACAGAGCAAGAGATTTGAACAAAGAGCGTGTACGAAAAA 660
DB 601 AAAGAAAAGTTCAGAACAGAGCAAGAGATTTGAACAAAGAGCGTGTACGAAAAA 660
QY 661 AATTTGAAAAGAAAAGAAAGCATGAGATATTTAGAGAGATTTATATGTCGTTTA 720
DB 661 AATTTGAAAAGAAAAGAAAGCATGAGATATTTAGAGAGATTTATATGTCGTTTA 720
QY 721 GAATACACAGCTATTAGAACTTCCATCAGAAAATGAAAGTGTATATATACCAATCAA 780
DB 721 GAATACACAGCTATTAGAACTTCCATCAGAAAATGAAAGTGTATATATACCAATCAA 780
QY 781 TCTTCTTTACCTCAGAGCAACAGAGGATATGAGATTTCCAGAAATATCTATATA 840
DB 781 TCTTCTTTACCTCAGAGCAACAGAGGATATGAGATTTCCAGAAATATCTATATA 840
QY 841 GAAAAAACAAATGAGAAATTTATTAACAAAGTTGGAAGAGAGGATATACATAAA 900
DB 841 GAAAAAACAAATGAGAAATTTATTAACAAAGTTGGAAGAGAGGATATACATAAA 900
QY 901 GGAATCTTGAAGAAAAGAAAGATGTTCAATAAAAACAGAACAAAAGAGATTAATCT 960
DB 901 GGAATCTTGAAGAAAAGAAAGATGTTCAATAAAAACAGAACAAAAGAGATTAATCT 960
QY 961 GGTGACATCAAAATCATCATTAAGAGACGTAATATTTCTGATGTTAATGATTTCAA 1020
DB 961 GGTGACATCAAAATCATCATTAAGAGACGTAATATTTCTGATGTTAATGATTTCAA 1020
QY 1021 ATTAAGTAAAGTAAAGATGAAATTAAGTCTGAATATGACGATTCATTAATGAAAGAA 1080
DB 1021 ATTAAGTAAAGTAAAGATGAAATTAAGTCTGAATATGACGATTCATTAATGAAAGAA 1080
QY 1081 GAAGATGATGAAGACTTGAAGCAATTTAAGCTTATGTCATATGCAATTTCCAAGAT 1140
DB 1081 GAAGATGATGAAGACTTGAAGCAATTTAAGCTTATGTCATATGCAATTTCCAAGAT 1140
QY 1141 GAAGAAAACATAGGAATTTATTAAGAACTAGAAAGTTGATAGAGAAAATGAAAATTTA 1200
DB 1141 GAAGAAAACATAGGAATTTATTAAGAACTAGAAAGTTGATAGAGAAAATGAAAATTTA 1200
QY 1201 GATGATTTAGATGAAGAAATGAAAAATCATCAGAGAAATTAATCTGAAGAAAATTA 1260
DB 1201 GATGATTTAGATGAAGAAATGAAAAATCATCAGAGAAATTAATCTGAAGAAAATTA 1260
QY 1261 AAAGGAAAAGAAATATGAAAAAAACAAAGATTAATTTTAAACCAATGATTAAGTTG 1320
DB 1261 AAAGGAAAAGAAATATGAAAAAAACAAAGATTAATTTTAAACCAATGATTAAGTTG 1320
QY 1321 TATGATGAGCATATTAATAAATATTAATAATGATTAAGCGATTAATAAGAAAAGAAA 1380
DB 1321 TATGATGAGCATATTAATAAATATTAATAATGATTAAGCGATTAATAAGAAAAGAAA 1380
QY 1381 TTCAATAAATCATTTGTTTCATATTTTGAAGAGAGCAATGAATTTTACAGATCGTGTAT 1440
DB 1381 TTCAATAAATCATTTGTTTCATATTTTGAAGAGAGCAATGAATTTTACAGATCGTGTAT 1440
QY 1441 GAGTTATCTGAAGATATTAATTAATTTATGAACTATATA 1482

Db 1441 GAGTATCTGAGATATACTAAATTTTATGAAACTATAA 1482

RESULT 5
US-08-098-327E-38
Sequence 38, Application US/0808327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22133-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-38

Query Match 97.6%; Score 1446; DB 3; Length 1493;
Best Local Similarity 99.8%; Pred. No. 2,9e-240;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1 CAAGAACAACAAGCATCTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAGAACAA 60
DB 1 CAAGAACAACAAGCATCTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAGAACAA 60
QY 61 CAAGCGATTTAGAACAGATAGACTTGTCTAAAGAAAGTTACAAAGACGCAAGCGAT 120
DB 61 CAAGCGATTTAGAACAGATAGACTTGTCTAAAGAAAGTTACAAAGACGCAAGCGAT 120
QY 121 TTAGAACAAGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAAGAGACTTGTCT -AAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 179
QY 181 GAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGT 240
DB 180 GAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGT 239
QY 241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTCTAAAGAA 300

DB 240 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTCTAAAGAA 299
QY 301 AAGTTACAAAGACGCAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAA 360
DB 300 AAGTTACAAAGACGCAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAA 359
QY 361 GAACAACAAGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAGAACAA 420
DB 360 GAACAACAAGCGATTTAGAACAAAGAGACGTCCT -AAGAAAGTTGCAAGAACAA 418
QY 421 AGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 480
DB 419 AGCGATTTAGAACAAAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 478
QY 481 GAACAAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAG 540
DB 479 GAACAAGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAG 538
QY 541 AGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGTCCT 600
DB 539 AGACGTCCT -AAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGTCCT 597
QY 601 AAAGAAAGTTGCAAGAGACGCAAGAGATTTAGAACAAAGAGGCTGATCGAAAAA 660
DB 598 AAAGAAAGTTGCAAGAGACGCAAGAGATTTAGAACAAAGAGGCTGATCGAAAAA 657
QY 661 AATTAGAAAGAAAGAAAGATGAGATATATTAGAGAGATTTATATGTCGTTA 720
DB 658 AATTAGAAAGAAAGAAAGATGAGATATATTAGAGAGATTTATATGTCGTTA 717
QY 721 GAATATACCAAGCTATAGAACTTCATCAGAAAAATGAACTGATATATATACATCAA 780
DB 718 GAATATACCAAGCTATAGAACTTCATCAGAAAAATGAACTGATATATATACATCAA 777
QY 781 TCTTCTTTTACTCAGAGAACAGAGGAAATAGAGATTCAGAGAAATATCTATATA 840
DB 778 TCTTCTTTTACTCAGAGAACAGAGGAAATAGAGATTCAGAGAAATATCTATATA 837
QY 841 GAAAAAACAATATAGAAATCTATTACAAACAATGTTGAAGAGCAAGGATATACATAA 900
DB 838 GAAAAAACAATATAGAAATCTATTACAAACAATGTTGAAGAGCAAGGATATACATAA 897
QY 901 GCACATCTTGAAGAAAGAAAGATGTTCAATTAACAAGCAACAAGAAAGATTAATCT 960
DB 898 GCACATCTTGAAGAAAGAAAGATGTTCAATTAACAAGCAACAAGAAAGATTAATCT 957
QY 961 GCTGACATACAAATCATACATTAGAGACGTAATATTTCTGATGTTATGATTTTCAA 1020
DB 958 GCTGACATACAAATCATACATTAGAGACGTAATATTTCTGATGTTATGATTTTCAA 1017
QY 1021 ATTAAGTATGAGAGATGAAGAAATAGTGTGAATATGACGATTCATTAATAGTAGAAG 1080
DB 1018 ATTAAGTATGAGAGATGAAGAAATAGTGTGAATATGACGATTCATTAATAGTAGAAG 1077
QY 1078 GAAGATGATGAAGATCTTAGAAGCAATTTAAAGCTATATGTCATATGACATTTCCAAAG 1137
QY 1141 GAAGAAACAATGGAATTTATTAAGAACTAGAGATTTGATAGAGAAAAATGAAAAATTA 1200
DB 1138 GAAGAAACAATGGAATTTATTAAGAACTAGAGATTTGATAGAGAAAAATGAAAAATTA 1197
QY 1201 GATGATTTAGAGAAAGAAATTAAGAAATCATAGAGAAATTTCTGAAGAAAAATTA 1260
DB 1198 GATGATTTAGAGAAAGAAATTAAGAAATCATAGAGAAATTTCTGAAGAAAAATTA 1257
QY 1261 AAAGAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTTTG 1320
DB 1258 AAAGAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATTAAGTTTG 1317
QY 1321 TATGATAGCATATTTAAAAATATTAAGATTAAGACGTTATATAGAAAAAGAAAA 1380

Db 1318 TATGATAGCATATTAAAAATATAAAAATGATAGAGGTTATATAGAAAAGAAAA 1377

Qy 1381 TTCATAAAATCATTTGTTTCATATATTTGACGAGACATGAAATTTTACAGATCGGAT 1440

Db 1378 TTCATAAAATCATTTGTTTCATATATTTGACGAGACATGAAATTTTACAGATCGGAT 1437

Qy 1441 GAGTATCTGAGATATATACCTAAATATTTTATGAACTATTA 1482

Db 1438 GAGTATCTGAGATATATACCTAAATATTTTATGAACTATTA 1479

RESULT 6

US-08-462-625-38

Sequence 38, Application US/08462625

Patent No. 6319502

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

APPLICANT: DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,625

FILING DATE: 05-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/098,327

FILING DATE: 24-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 91 01286

FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 1493 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

US-08-462-625-38

Query Match 97.6%; Score 1446; DB 3; Length 1493;

Best Local Similarity 99.8%; Pred. No. 2,9e-240;

Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Qy 1 CAAGAACAAAGGATCTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 60

Db 1 CAAGAACAAAGGATCTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 60

Qy 61 CAAGCGGTTTGAACAAAGATGAGCTTGCTAAAGAAAAGTTTCAAGAGCAGCAAGCGAT 120

Db 61 CAAGCGATTTAGAACAAAGATAGACTTGCTTAAAGAAAAGTTTCAAGAGCAGCAAGCGAT 120

Qy 121 TTGCAACAGAGAGACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTTAGAACAA 180

Db 121 TTGCAACAGAGAGACTTGCT - AAGAAAAGTTGCAAGAACAAAGCGATCTTAGAACAA 179

Qy 181 GAGGAGCGTCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCT 240

Db 180 GAGGAGCGTCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCT 239

Qy 241 GCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATAGACTTGCTTAAAGAA 300

Db 240 GCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATAGACTTGCTTAAAGAA 299

Qy 301 AAGTTACAGAGCAGCAAGCGATTTAGAACAGAGAGAGCTGCTTAAAGAAAAGTTGCCAA 360

Db 300 AAGTTACAGAGCAGCAAGCGATTTAGAACAGAGAGAGCTGCTTAAAGAAAAGTTGCCAA 359

Qy 361 GAACAAACAAAGCGATTTAGAACAGAGAGAGCTGCTTAAAGAAAAGTTGCCAAAGAACAA 420

Db 360 GAACAAACAAAGCGATTTAGAACAGAGAGAGCTGCT - AAGAAAAGTTGCCAAAGAACAA 418

Qy 421 AGCGATTTAGAACAGAGAGACTTGCTTAAAGAAAAGTTGCCAAAGAACAAAGCGATTTA 480

Db 419 AGCGATTTAGAACAGAGAGACTTGCTTAAAGAAAAGTTGCCAAAGAACAAAGCGATTTA 478

Qy 481 GAACAGAGAGAGCTGCTTAAAGAAAAGTTGCCAAAGAACAAAGCGATTTTGAACAGAG 540

Db 479 GAACAGAGAGAGCTGCTTAAAGAAAAGTTGCCAAAGAACAAAGCGATTTTGAACAGAG 538

Qy 541 AGAGCTGCTTAAAGAAAAGTTGCCAAAGAACAAAGCGATTTTGAACAGAGAGAGCTGCT 600

Db 539 AGAGCTGCT - AAGAAAAGTTGCCAAAGAACAAAGCGATTTTGAACAGAGAGAGCTGCT 597

Qy 601 AAGAAAAGTTGCCAAAGAGCAGCAAGAGATTTAGAACAAAGAGCTGATACGAAAAAA 660

Db 598 AAGAAAAGTTGCCAAAGAGCAGCAAGAGATTTAGAACAAAGAGCTGATACGAAAAAA 657

Qy 661 AATTAGAAAAGAAAAGAGAAATGAGATATATTTAGAGAGATTTATATGTCGTTTA 720

Db 658 AATTAGAAAAGAAAAGAGAAATGAGATATATTTAGAGAGATTTATATGTCGTTTA 717

Qy 721 GAAATACAGCTATAGAACTTCCATCAGAAAAGAAAGTGGATTTATATACCATCA 780

Db 718 GAAATACAGCTATAGAACTTCCATCAGAAAAGAAAGTGGATTTATATACCATCA 777

Qy 781 TCTTCTTTACCTCAGAGCAACAGAGGATAGAGATTTCCAGAAAATATCTATATATA 840

Db 778 TCTTCTTTACCTCAGAGCAACAGAGGATAGAGATTTCCAGAAAATATCTATATATA 837

Qy 841 GAAAAAACAATAGAGAAATCTATTTACAAATGTTGAGAGAGAGGATATACATATA 900

Db 838 GAAAAAACAATAGAGAAATCTATTTACAAATGTTGAGAGAGAGGATATACATATA 897

Qy 901 GGAATCTTGAAGAAAAGAAAGAGTGTTCATATAAACAGAACAAAAAGAGATTAATCT 960

Db 898 GGAATCTTGAAGAAAAGAAAGAGTGTTCATATAAACAGAACAAAAAGAGATTAATCT 957

Qy 961 GCTGACATACAAAATCATATCTTAGAGACAGTAAATTTCTGATGTTAATGATTTCAA 1020

Db 958 GCTGACATACAAAATCATATCTTAGAGACAGTAAATTTCTGATGTTAATGATTTCAA 1017

Qy 1021 ATAAATAGATAGAGATGAGATTAAGTCTGAATATGACGATTTCAATTAATGATGAAGAA 1080

Db 1018 ATAAATAGATAGAGATGAGATTAAGTCTGAATATGACGATTTCAATTAATGATGAAGAA 1077

Qy 1081 GAAGATAGATAGAGCTTAGAGCAATTTTAAGCTTATGCAATATGCAATTTCCAGAT 1140

Db 1078 GAAGATAGATAGAGCTTAGAGCAATTTTAAGCTTATGCAATATGCAATTTCCAGAT 1137

Qy 1141 GAAGAAAACATAGGAATTTTAAAGAACTAGAGATTTGATAGAGAAAATGAAAATTTTA 1200

Db 1138 GAAGAAAACATAGGAATTTTAAAGAACTAGAGATTTGATAGAGAAAATGAAAATTTTA 1197

APPLICANT: DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22133-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-32

Query Match 64.1%; Score 950; DB 3; Length 950;
Best Local Similarity 100.0%; Pred. No. 5.8e-155;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

11 AAGCGATCTGAAACAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 70
1 AAGCGATCTGAAACAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 60
71 TAGAACAAGATAGACTTCTTAAGAAAAGTTCAAGAGCGAAAGCGATTAGAACAA 130
61 TAGAACAAGATAGACTTCTTAAGAAAAGTTCAAGAGCGAAAGCGATTAGAACAA 120
131 AGAGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAGAGAGCGTG 190
121 AGAGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAGAGAGCGTG 180
191 CTAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAGAGAGCGTCTTAAGAAA 250
181 CTAAGAAAAGTTGCAAGAACAAAGCGATCTTAAGAACAGAGAGCGTCTTAAGAAA 240
251 AGTTGCAAGAACAAAGCGATCTTAAGAACAGATAGACTTCTTAAGAAAAGTTA 310
241 AGTTGCAAGAACAAAGCGATCTTAAGAACAGATAGACTTCTTAAGAAAAGTTA 300
311 AGAGCAAGCGATTTAGAACAAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAA 370
301 AGAGCAAGCGATTTAGAACAAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAA 360

371 GCGATTTAGAACAAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTAG 430
361 GCGATTTAGAACAAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTAG 420
431 AACAGAGAGACTTGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 490
421 AACAGAGAGACTTGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 480
491 GACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTA 550
481 GACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTA 540
551 AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTTAAGAAAAGT 610
541 AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTTAAGAAAAGT 600
611 TGCAGAGCGCAAAAGATTTAGAACAAAGAGCGTGATACGAAAAAAATTTAGAAA 670
601 TGCAGAGCGCAAAAGATTTAGAACAAAGAGCGTGATACGAAAAAAATTTAGAAA 660
671 GAAAAAAGAACATGAGATATATTAGACAGAGATTTATATGCTTTAGAAAATACAG 730
661 GAAAAAAGAACATGAGATATATTAGACAGAGATTTATATGCTTTAGAAAATACAG 720
731 CTATAGACTTCCATCAGAAAATGAAACGTGATTTATATACACATCATCTTCTTAC 790
721 CTATAGACTTCCATCAGAAAATGAAACGTGATTTATATACACATCATCTTCTTAC 780
791 CTCAGAGCAACAGAGGAATAGTAGAGATTCCAAGGAATATCTATATAGAAAAACAA 850
781 CTCAGAGCAACAGAGGAATAGTAGAGATTCCAAGGAATATCTATATAGAAAAACAA 840
851 ATAGAGATCTATTACACAAATGTTGAAGAGCAAGAGATATATAGAAAGCATCTTG 910
841 ATAGAGATCTATTACACAAATGTTGAAGAGCAAGAGATATATAGAAAGCATCTTG 900
911 AAGAAAAGAAAAGTGGTTCAATTAACACAGAACAAAAAGAAAGTAATCT 960
901 AAGAAAAGAAAAGTGGTTCAATTAACACAGAACAAAAAGAAAGTAATCT 950

RESULT 9
US-08-257-073-15
Sequence 15, Application US/08257073
Patent No. 5766597
GENERAL INFORMATION:
APPLICANT: Paolietti, Enzo
APPLICANT: de Taisne, Charles
TITLE OF INVENTION: MALARIA RECOMBINANT POXYVIRUS VACCINE
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P. C.
STREET: 530 Fifth Avenue, 25th Floor
CITY: New York
STATE: New York
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/257,073
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/852,305

FILING DATE: 18-MAR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/672,183
 FILING DATE: 20-MAR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Prommer, William S.
 REGISTRATION NUMBER: 25,506
 REFERENCE/DOCKET NUMBER: 454310-2570
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 840-3333
 TELEFAX: (212) 840-0712
 TELEX: 425066 CURTMS
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1320 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-257-073-15

Query Match 57.6%; Score 853.2; DB 2; Length 1320;
 Best Local Similarity 87.9%; Pred. No. 2,7e-138;
 Matches 930; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

425 ATTGAGAACAGAGAGACTTGTCTTAAAGAAAGTTGCAAGACAAAGCCATTAGAAC 484
 263 ATGTAAAAATGTGTACCAACAAATTTCAAAAGTCTTTTAAAGAAATCTTGTTCTCG 322
 485 AAGAGAGAGCTGCTTAAAGAAAGTTGCAAGACAAAGCCATTAGAACAGAGAC 544
 323 AGAATATATCTCTTAAAGAAAGTTTAAATAGAGAAAGAAATTAATGAAACATTA 382
 545 GTGCTTAAAGAAAGTTGCAAGACAAAGCCATTAGAACAGAGAGCTGCTTAAG 604
 383 TAAATGATGATGACGATTAATAAAATATATTTAAAGGCGAAGCAAGAAACAGAC 442
 605 AAAAGTTGCAAGAGCGCAAGAGATTGAAACAAAGAAAGGCTGTAAGAAAAAAT 664
 443 ATCTTGAAGAAAAAGCGCGCGCTTAAAGAAACGAGAAAGGCTGTAAGAAAAAAT 502
 665 TAGAAAGAAAAAGAAACATGAGATATATTAGCAGAGATTTATATGCTGTTAGAA 724
 503 TAGAAAGAAAAAGAAACATGAGATATATTAGCAGAGATTTATATGCTGTTAGAA 562
 725 TACACCTATAGAACTTCCATCAGAAATGAAAGCTGATATTATATCAATCAATCTT 784
 563 TACACCTATAGAACTTCCATCAGAAATGAAAGCTGATATTATATCAATCAATCTT 622
 785 CTTTACCTCAGGCAACAGAGGAAATGATGAGATTTCAAAGAAATATCTAATAGAA 844
 623 CTTTACCTCAGGCAACAGAGGAAATGATGAGATTTCAAAGAAATATCTAATAGAA 682
 845 AAAACAATAGAGATCTATTTACAAACAATGTTGAAGAGCAAGAGATATATCAATAGAG 904
 683 AAAACAATAGAGATCTATTTACAAACAATGTTGAAGAGCAAGAGATATATCAATAGAG 742
 905 ATCTTGAAGAAAAAGAAAGATGTTCAATTAACCAAGACAAAGAAAGATTAATCTGCTG 964
 743 ATCTTGAAGAAAAAGAAAGATGTTCAATTAACCAAGACAAAGAAAGATTAATCTGCTG 802
 965 ACCTACAAATCTATCTTATAGACAGCTAAATTTCTGATGTTATGATTTCAATTA 1024
 803 ACCTACAAATCTATCTTATAGACAGCTAAATTTCTGATGTTATGATTTCAATTA 862
 1025 GTAAATGATGAGATGAATTAAGTGTGATATGATGATGATGATGATGATGATGATG 1084
 863 GTAAATGATGAGATGAATTAAGTGTGATATGATGATGATGATGATGATGATGATG 922
 1085 ATGATGAAGATTAAGAGCAATTTAAGCTTATTTGCAATTAAGCAATTTCAAGATGAAG 1144
 923 ATGATGAAGATTAAGAGCAATTTAAGCTTATTTGCAATTAAGCAATTTCAAGATGAAG 982
 1145 AAAACAATAGAGATTTATTAAGAAACATGAGATTTGATGAGAAATTAATTAATGATG 1204

983 AAAACAATAGAGATTTATTAAGAACTAGAAAGATTTGATGAGAAAAATCAAAATTTAGATG 1042
 1205 ATTTAGATGAAGAAATGAATAATCATCAGAGAAATTTCTGAGAAAAATTAAGAAAG 1264
 1043 ATTTAGATGAAGAAATGAATAATCATCAGAGAAATTTCTGAGAAAAATTAAGAAAG 1102
 1265 GAAAGAAATATGAAGAAAAAGCAAGATTAATTTTAAACCAATGATTAAGATTTGATG 1324
 1103 GAAAGAAATATGAAGAAAAAGCAAGATTAATTTTAAACCAATGATTAAGATTTGATG 1162
 1325 ATGAGCATATTTAAATAATATATAATGATTAAGCGGTTAATTAAGAAAAAGAAATTTCA 1384
 1163 ATGAGCATATTTAAATAATATATAATGATTAAGCGGTTAATTAAGAAAAAGAAATTTCA 1222
 1385 TAAATCATTTCTTATCATATTTGACGAGACAAATGAATTTTACAGATCGATGAGT 1444
 1223 TAAATCATTTCTTATCATATTTGACGAGACAAATGAATTTTACAGATCGATGAGT 1282
 1445 TATCTGAAGATATACTAATATTTTATGAACATTA 1482
 1283 TATCTGAAGATATACTAATATTTTATGAACATTA 1320

RESULT 10
 US-08-098-327E-34
 ; Sequence 34; Application US/08098327E
 ; Patent No. 6270771
 ; GENERAL INFORMATION:
 ; APPLICANT: GUERIN-MARCHAND, Claudine
 ; APPLICANT: DRULIHE, Pierre
 ; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
 ; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
 ; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
 ; NUMBER OF SEQUENCES: 46
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: P. O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/098,327E
 ; FILING DATE: 24-NOV-1993
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 91 01286
 ; FILING DATE: 05-FEB-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McGowan, Malcolm K.
 ; REGISTRATION NUMBER: 39,300
 ; REFERENCE/DOCKET NUMBER: 010830-045
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 836-6620
 ; TELEFAX: (703) 836-2021
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 988 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; PUBLICATION INFORMATION:
 ; DOCUMENT NUMBER: WO 92/13884
 ; PUBLICATION DATE: 20-AUG-1992
 US-08-098-327E-34

Query Match 31.9%; Score 472.8; DB 3; Length 988;
Best Local Similarity 95.7%; Pred. No. 7.3e-73;
Matches 486; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

77 AAGATAGACTTGTCTAAAGAAAAGTTACAGAGCAGCAAGGCGATTTAGAAACAAGAGAGAC 136
|||
481 AAGAAAAAGCAGCTTAAGAAAAGTTACAGGCGCAACAAAGCGATTCAGAACAGAGAGAC 540
|||
137 TTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTAGAACAGAGAGAGAGCGTCTAAAG 196
|||
541 GTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTAGAACAGAGAGAGAGCGTCTAAAG 600
|||
197 AAAAGTTGCAAGAACCAACAAAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGC 256
|||
601 AAAAGTTGCAAGAACCAACAAAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGC 660
|||
257 AAGAACCAACAAAGGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTAGAACAGAGAGC 316
|||
661 AAGAACCAACAAAGGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTAGAACAGAGAGC 720
|||
317 AAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACCAACGCGATT 376
|||
721 AAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACCAACGCGATT 780
|||
377 TAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAG 436
|||
781 TAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAG 840
|||
437 AGAGACTTGTCTAAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAGAGAGAGCGTG 496
|||
841 AGAGACTTGTCTAAAGAAAAGTTGCAAGAGCAGCAAGGCGATTTAGAACAGATGACTTG 900
|||
497 CTAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAGAGAGAGCGTCTTAAGAAA 556
|||
901 CTAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAGAGAGAGCGTCTTAAGAAA 960
|||
557 AGTTGCAAGAACCAACAAAGCGATTTAGA 584
|||
961 GGTTCGAAGAACCAACAAAGCGATTTAGA 988
|||

RESULT 11
US-08-462-625-34
; Sequence 34, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,625
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/098,327
; FILING DATE: 24-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286

FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 988 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
US-08-462-625-34

Query Match 31.9%; Score 472.8; DB 3; Length 988;
Best Local Similarity 95.7%; Pred. No. 7.3e-73;
Matches 486; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

77 AAGATAGACTTGTCTAAAGAAAAGTTACAGAGCAGCAAGGCGATTTAGAAACAAGAGAGAC 136
|||
481 AAGAAAAAGCAGCTTAAGAAAAGTTACAGGCGCAACAAAGCGATTCAGAACAGAGAGAC 540
|||
137 TTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTAGAACAGAGAGAGAGCGTCTAAAG 196
|||
541 GTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTAGAACAGAGAGAGAGCGTCTAAAG 600
|||
197 AAAAGTTGCAAGAACCAACAAAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGC 256
|||
601 AAAAGTTGCAAGAACCAACAAAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGC 660
|||
257 AAGAACCAACAAAGGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTAGAACAGAGAGC 316
|||
661 AAGAACCAACAAAGGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTAGAACAGAGAGC 720
|||
317 AAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACCAACGCGATT 376
|||
721 AAGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACCAACGCGATT 780
|||
377 TAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAG 436
|||
781 TAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAG 840
|||
437 AGAGACTTGTCTAAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAGAGAGAGCGTG 496
|||
841 AGAGACTTGTCTAAAGAAAAGTTGCAAGAGCAGCAAGGCGATTTAGAACAGATGACTTG 900
|||
497 CTAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAGAGAGAGCGTCTTAAGAAA 556
|||
901 CTAAGAAAAGTTGCAAGAACCAACGCGATTTAGAACAGAGAGAGCGTCTTAAGAAA 960
|||
557 AGTTGCAAGAACCAACAAAGCGATTTAGA 584
|||
961 GGTTCGAAGAACCAACAAAGCGATTTAGA 988
|||

RESULT 12
US-08-098-3278-37
; Sequence 37, Application US/080983278
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:

Db 509 GTGCTAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGACTTGCTAAAG 568
Qy 197 AAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGC 256
Db 569 AAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGC 628
Qy 257 AAGAACAAACAAGGAGTTTGAACAAGATGACTTGTAAAGAAAAGTTTCAAGAGAGC 316
Db 629 AAGAACAAACAAGGAGTTTGAACAAGAGACTTGTAAAGAAAAGTTTCAAGAACAC 688
Qy 317 AAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 376
Db 689 AAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATT 748
Qy 377 TAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAG 436
Db 749 TAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAG 808
Qy 437 AGAGACTTGTCAAGAACAGCGATTAGAACAAAGCGATTAGAACAGAGAGAGCTG 496
Db 809 AGAGACTTGTCAAGAACAGCGATTAGAACAAAGCGATTAGAACAGAGAGAGCTG 868
Qy 497 CTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGCGTCTAAAGAAA 556
Db 869 CTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGCGTCTAAAGAAA 928
Qy 557 AGTTGCAAGAACAAAGCGATT 582
Db 929 GGTTGCAAGAACAAAGCGATT 954

RESULT 14

US-09-902-540-1280
; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(1039)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

Query Match 8.8%; Score 129.8; DB 3; Length 1039;
Best Local Similarity 47.2%; Pred. No. 7.4e-14;
Matches 493; Conservative 0; Mismatches 544; Indels 8; Gaps 3;

Qy 306 ACAAGACGACGAACGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAAGAGACA 365
Db 2 AAATTAAGCTTC 61
Qy 366 ACAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAAGAGACAAAGCGA 425
Db 62 AAATTAAGCTTC 121
Qy 426 TTAGAACAGAGAGACTTGTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAGACA 485
Db 122 TAAAAAAAAAATTAACAATTAATAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAATA 181

Qy 486 AGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGC 545
Db 182 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAATA 241
Qy 546 TGCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGCGTCTAAAGA 605
Db 242 TATTAATAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAATA 301
Qy 606 AAAGTTGCAAGAGAGAGAGATTAGAACAAAGAGAGCGTCTAATCGAAAAAATTT 665
Db 302 AAAAAAAAAAATTAATAAAAAAAAAAATTAATAATAATAATAATAATAATAATAAT 361
Qy 666 AGAAGAAAAGAGACATGAGATTAATTAAGAGAGATTATATGCTGTTAGAAAT 725
Db 362 AACAAAAAACAATAAATTAATAATAATAATAATAATAATAATAATAATAATAAAT 421
Qy 726 ACCAGATTAAGACTTCATCAGAAAATGACGTGATTAATATACATCATGATTC 785
Db 422 AAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAACAACA-AAAAACCAAAAAAAAA 480
Qy 786 TTTACTCAGACACAGAGGAAATGTAGAGATTCAGAGAAATATCTAATAGAAA 845
Db 481 ATAAATTAATAAAAAAAAAAATGAACAATAATAATAATAATAATAATAATAATA 540
Qy 846 AACAAATAGAGATCTATTACAATAATGTTGAAGAGCGAAGGATTAATATAAGACA 905
Db 541 AACAAATTAATAAAAAAAAAAATGAATAAAAAAAAAAAGTAATAATTAATTTA 600
Qy 906 TCTTGAAGAAAAGAGATGTTCAATTAACCAAGAACAAAGAGATTAATCTGCTGA 965
Db 601 AAAAAAATAAAAAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAA 660
Qy 966 CATACAAATCATACATTAGAGACAGTAATATTTCTGATGTTAATGATTTCAATAG 1025
Db 661 AAAAAAGTTTCTTAATAAAAAAAAAAATCTCCCTAATAACAATA-AAAAA 714
Qy 1026 TAAGTATAGAGATGAATAATAGTCTGATATGACGATTCATTAATAGATGAAGAGAGA 1085
Db 715 AAAAAAAGAAATTAACAAAGCTTAATCAACCCCAATATTAACAAATAATCCTAATA 774
Qy 1086 TGATGAAGACTTGAAGCAATTTTAAGCTATTTGCAATATGACAAATTTCCAGATGAGA 1145
Db 775 CAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 834
Qy 1146 AAACATAGAAATTTAT-AAAAGACTAGAGATTTGATAGAAATAAGAAATTTAGATG 1204
Db 835 AAATGAGGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 894
Qy 1205 ATTTAGATGAAGATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1264
Db 895 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 954
Qy 1265 GAAAGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1324
Db 955 AAAAAAAAAAAAAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATA 1014
Qy 1325 ATGAGCATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1349
Db 1015 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1039

RESULT 15

US-09-902-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540

;/ CURRENT FILING DATE: 2001-07-10
;/ PRIOR APPLICATION NUMBER: 60/217,883
;/ PRIOR FILING DATE: 2000-07-10
;/ NUMBER OF SEQ ID NOS: 16825
;/ SEQ ID NO: 1357
;/ LENGTH: 612
;/ TYPE: DNA
;/ ORGANISM: Myxococcus xanthus
;/ FEATURE:
;/ NAME/KEY: unsure
;/ LOCATION: (1)..(612)
;/ OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 8.2%; Score 120.8; DB 3; Length 612;
Best Local Similarity 50.3%; Pred. No. 2.5e-12;
Matches 293; Conservative 0; Mismatches 289; Indels 0; Gaps 0;

Qy	808	AATGATGAGATTCCAAAGAAATATCTAATATAGAAAAACAATAGAAATCTATTACA	867
Db	9	AA	68
Qy	868	ACAAATGTTGAAGACGAAGGATATACATAAGACATCTTGAGAAAAAGAAATGCT	927
Db	69	AA	128
Qy	928	TCATATTAACCAACAAAGAAAGATTAATCTGCTGACATCAAAATCATACATTAGAG	987
Db	129	AA	188
Qy	988	ACAGTAAATATTTCTGATGATTTTCAAAATAGTAGTAGAGATGAATTAAGT	1047
Db	189	AA	248
Qy	1048	GCTGAATATGACGATTCATTATAGATGAGAAAGAAAGATGAGACTTAGACGAATT	1107
Db	249	AA	308
Qy	1108	AAGCCATTTGTCATATATGACAAATTTCCAAAGTGAAGAAACATNGAATTTATTAAGAA	1167
Db	309	AAANATTAATTAATAAAAAAAAAAAAAAAAAAATTAATAATAATAATAATAATAA	368
Qy	1168	CTAGAAGATTTGTAGAGAAAAATGAATAATTTAGATGATTTAGATGAAGAAATAGAAAA	1227
Db	369	ATTAATAAAAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAA	428
Qy	1228	TCATCAGAAAGATTATCTGAGAAAAATTAATAATAATAATAATAATAATAATAATAA	1287
Db	429	AAAAAAAAAAAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAA	488
Qy	1288	GATTAATTAATTTTAAACCAATGATTAATAAGTTTGTATGATGAGCATTTAATAATAA	1347
Db	489	AA	548
Qy	1348	AATGATAGAGGTTAATAGAAAAAGAAAAATTCATAAAA	1389
Db	549	AAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	590

Search completed: December 7, 2005, 03:40:57
Job time : 302 secs

PT vaccination against, treatment of and diagnosis of malaria.
XX
PS Claim 2; Fig 8-10; 81pp; French.

XX This nucleotide sequence is the 3' part of the P.falciparum liver-stage
CC specific antigen (LSA) gene. It codes for a polypeptide sequence which
CC carries a T cell epitope characteristic of a protein produced in
CC hepatocytes infected with P.falciparum. The polypeptide can be used in
CC the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX

SQ Sequence 1496 BP; 716 A; 169 C; 300 G; 311 T; 0 U; 0 Other;

Query Match 100.0%; Score 1482; DB 2; Length 1496;

Best Local Similarity 100.0%; Pred. No. 8.5e-202; Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 CAAGAACAAACAAGCGATCTAGAACAGAGACGCTCTAAAGAAAAGTTGCAGAACAA 60
Db 1 CAAGAACAAACAAGCGATCTAGAACAGAGACGCTCTAAAGAAAAGTTGCAGAACAA 60
Oy 61 CAAAGCGATTTAGAACAGATGACTTGTCTAAAGAAAAGTTGCAGAGCAGCAAGCGAT 120
Db 61 CAAAGCGATTTAGAACAGATGACTTGTCTAAAGAAAAGTTGCAGAGCAGCAAGCGAT 120
Oy 121 TTAGAACAAAGAGACTTCTAAAGAAAAGTTGCAGAACAAACAAGCGATCTAGAACAA 180
Db 121 TTAGAACAAAGAGACTTCTAAAGAAAAGTTGCAGAACAAACAAGCGATCTAGAACAA 180
Oy 181 GAGAGCGTCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTAGAACAGAGAGACGT 240
Db 181 GAGAGCGTCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTAGAACAGAGAGACGT 240
Oy 241 GCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTAGAACAGATGACTTGTCTAAAGAA 300
Db 241 GCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTAGAACAGATGACTTGTCTAAAGAA 300
Oy 301 AAGTTACAAGAGAGAGCAAGCGATTTAGAACAGAGAGACGTCTAAAGAAAAGTTGCA 360
Db 301 AAGTTACAAGAGAGCAAGCGATTTAGAACAGAGAGACGTCTAAAGAAAAGTTGCA 360
Oy 361 GAACAACAAGCGATTTAGAACAGAGAGACGTCTAAAGAAAAGTTGCAGAACAAACA 420
Db 361 GAACAACAAGCGATTTAGAACAGAGAGACGTCTAAAGAAAAGTTGCAGAACAAACA 420
Oy 421 ACGGATTTAGAACAGAGAGACCTTCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTA 480
Db 421 ACGGATTTAGAACAGAGAGACCTTCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTA 480
Oy 481 GAACAAGAGAGACGTCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTAGAACAGAG 540
Db 481 GAACAAGAGAGACGTCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTAGAACAGAG 540
Oy 541 AGACGTCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTAGAACAGAGAGACGTCT 600
Db 541 AGACGTCTAAAGAAAAGTTGCAGAACAAACAAGCGATTTAGAACAGAGAGACGTCT 600
Oy 601 AAAGAAAAGTTGCAGAGCAGCAAGAGATTTAGAACAAAGGAGCGTATAGAAAAAA 660
Db 601 AAAGAAAAGTTGCAGAGCAGCAAGAGATTTAGAACAAAGGAGCGTATAGAAAAAA 660
Oy 661 AATTTGAAAAGAAAAGAGAACTGAGATATTTTGCAGAGATTTATATGTCGTTTA 720
Db 661 AATTTGAAAAGAAAAGAGAACTGAGATATTTTGCAGAGATTTATATGTCGTTTA 720
Oy 721 GAAATACCAAGCTTAGAACTTCATCAGAAATGAACTGAGATATTTATACCACTCAA 780
Db 721 GAAATACCAAGCTTAGAACTTCATCAGAAATGAACTGAGATATTTATACCACTCAA 780
Oy 781 TCTTCTTTAAGCTCAGAGCAACAGAGGAAATAGTAGAGATTCAGAGAAATATCTATATA 840
Db 781 TCTTCTTTAAGCTCAGAGCAACAGAGGAAATAGTAGAGATTCAGAGAAATATCTATATA 840
```

```

Oy 841 GAAAAAACAAATAGGAATCTTATTACAACAATGTTGAGAGCAGAGGATATACATATA 900
Db 841 GAAAAAACAAATAGGAATCTTATTACAACAATGTTGAGAGCAGAGGATATACATATA 900
Oy 901 GGACATCTTTGAGAAAAGAGATGCTTCAATTAAGCCAGAACAAAGAAATGAAATCT 960
Db 901 GGACATCTTTGAGAAAAGAGATGCTTCAATTAAGCCAGAACAAAGAAATGAAATCT 960
Oy 961 GCTGACATACAAATCATACATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCA 1020
Db 961 GCTGACATACAAATCATACATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCA 1020
Oy 1021 ATTAGTAATGATGAGATGAAATAGTGTCTGAATATGACATTCATTAATAGTAAAGA 1080
Db 1021 ATTAGTAATGATGAGATGAAATAGTGTCTGAATATGACATTCATTAATAGTAAAGA 1080
Oy 1081 GAAAGTATGAAAGACTTGAAGCAATTTAAGCCTATGTGCAATATGACAAATTTCCAAGAT 1140
Db 1081 GAAAGTATGAAAGACTTGAAGCAATTTAAGCCTATGTGCAATATGACAAATTTCCAAGAT 1140
Oy 1141 GAAGAAAACATAGCAATTTATTAAGACCTAGAAAGTTGATGAGAAAAATGAAATTTTA 1200
Db 1141 GAAGAAAACATAGCAATTTATTAAGACCTAGAAAGTTGATGAGAAAAATGAAATTTTA 1200
Oy 1201 GATGATTTAGATGAAGAAATGAAAAATCATCAGAAATTTATCTGAAGAAAAATTAATA 1260
Db 1201 GATGATTTAGATGAAGAAATGAAAAATCATCAGAAATTTATCTGAAGAAAAATTAATA 1260
Oy 1261 AAAGAAAAGAAATATGAAAAAAACAAAGATTAATTTTAAACCAAGATTAAGTTTG 1320
Db 1261 AAAGAAAAGAAATATGAAAAAAACAAAGATTAATTTTAAACCAAGATTAAGTTTG 1320
Oy 1321 TATGATGACATATTAATAAATATTAATAATGTAAGCAGTTATTAAGAAAAAGAAAA 1380
Db 1321 TATGATGACATATTAATAAATATTAATAATGTAAGCAGTTATTAAGAAAAAGAAAA 1380
Oy 1381 TTCAATTAATCATCTGTTTCATATATTTTGAACGAGCAATGAATTTTACAGATCGTGAT 1440
Db 1381 TTCAATTAATCATCTGTTTCATATATTTTGAACGAGCAATGAATTTTACAGATCGTGAT 1440
Oy 1441 GAGTTATCTGAAGATATTAATTAATTTTATGAACCTATTA 1482
Db 1441 GAGTTATCTGAAGATATTAATTAATTTTATGAACCTATTA 1482

RESULT 2
AAQ28115
ID AAQ28115 standard; DNA; 950 BP.
XX
AC AAQ28115;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum LSA-R-NR coding sequence.
XX
KW Malaria; hepatocyte; sporozoite; Plasmodium DG 536; T-cell epitope;
KW paludism; liver stage-specific antigen; ss.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FH repeat_region 3..629 /*tag= a
FH FT /*tag= a TANDEM
FH FT repeat_unit 3..53 /*tag= b
FH FT
XX PN MO9213884-A1.
XX XX 20-AUG-1992.
XX XX 05-FEB-1992; 92WO-FR000104.
```

XX 05-FEB-1991; 91FR-00001286.
XX (INSP) INST PASTEUR.
XX
PI Querimarchand C, Drulhe P;
XX WPI, 1992-299985/36.
XX P-PSDB; AAR26941.
XX
PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
PS Claim 17, Fig 2, 81pp; French.
XX
XX A genomic DNA bank of *P. falciparum* EcoRI fragments prepared in lambda
CC gII1 was used to transform *E. coli*. The expression library was screened
CC with human anti-liver stage antigens of all stages of *P. falciparum*. The
CC library was rescreened with antibodies affinity-purified on a clone
CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic LSA
CC epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
CC containing a 12-repeat region followed by a non-repeat region. Preferred
CC antigenic polypeptides of the invention are derived from the amino acid
CC sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 950 BP; 456 A; 127 C; 206 G; 161 T; 0 U; 0 Other;
Query Match 64.1%; Score 950; DB 2; Length 950;
Best Local Similarity 100.0%; Pred. No. 2,9e-126;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAAGCGATCTAGAACACAGAGAGCGTCTAAAGAAAGTTGCAAGAACACAAAGCGATT 70
DB 1 AAAGCGATCTAGAACACAGAGAGCGTCTAAAGAAAGTTGCAAGAACACAAAGCGATT 60
QY 71 TAGAACACAGATGACTTCTTAAGAAAGTTACAAAGACACAAAGCGATTTAAGACAG 130
DB 61 TAGAACACAGATGACTTCTTAAGAAAGTTACAAAGACACAAAGCGATTTAAGACAG 120
QY 131 AGAGACTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACACAGAGAGCGTG 190
DB 121 AGAGACTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACACAGAGAGCGTG 180
QY 191 CTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACACAGAGAGCGTGCTTAAGAA 250
DB 181 CTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACACAGAGAGCGTGCTTAAGAA 240
QY 251 AGTTGCAAGAACACAAAGCGATTTAGAACAGATGACTTCTTAAGAAAGTTACAG 310
DB 241 AGTTGCAAGAACACAAAGCGATTTAGAACAGATGACTTCTTAAGAAAGTTACAG 300
QY 311 AGCAGCAAGGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACACAA 370
DB 301 AGCAGCAAGGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACACAA 360
QY 371 GCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACACAAAGCGATTAG 430
DB 361 GCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAACACAAAGCGATTAG 420
QY 431 AACACAGAGAGACTGCTTAAGAAAGTTGCAAGAACACAAAGCGATTTAGACAGAG 490
DB 421 AACACAGAGAGACTGCTTAAGAAAGTTGCAAGAACACAAAGCGATTTAGACAGAG 480
QY 491 GACGCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTA 550
DB 481 GACGCTTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGTGCTA 540
QY 551 AAGAAAGTTGCAAGAACACAAAGCGATTTAGAACAGAGAGCGTGCTTAAGAAAGT 610
DB 541 AAGAAAGTTGCAAGAACACAAAGCGATTTAGAACAGAGAGCGTGCTTAAGAAAGT 600
QY 611 TGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCGTGATTCGAAAAAAATTTAGAA 670

DB 601 TGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCGTGATACGAAAAAAATTTAGAAA 660
QY 671 GAAAAAGGAAACATGAGATATATTAGCAGAGATTTATATGCTGTTTGAATACAG 730
DB 661 GAAAAAGGAAACATGAGATATATTAGCAGAGATTTATATGCTGTTTGAATACAG 720
QY 731 CTATAGAACTTCATCGAAGAAATGAACGTGATATATTATACACATCAATCTCTTAC 790
DB 721 CTATAGAACTTCATCGAAGAAATGAACGTGATATATTATACACATCAATCTCTTAC 780
QY 791 CTCAGACAAACAGAGGAAATAGTAGATTCCAGAAATATCTATATAGAAAAACAA 850
DB 781 CTCAGACAAACAGAGGAAATAGTAGATTCCAGAAATATCTATATAGAAAAACAA 840
QY 851 ATAGAGATCTATTACAAACAAATGTTGAGAGCAGAAAGGATATATACAAAGACATTTG 910
DB 841 ATAGAGATCTATTACAAACAAATGTTGAGAGCAGAAAGGATATATCAATAAAGCATCTTG 900
QY 911 AAGAAAAAGAAATGCTTCAATTAACCAAGCAAAAAAGAAATCT 960
DB 901 AAGAAAAAGAAATGCTTCAATTAACCAAGCAAAAAAGAAATCT 950
RESULT 3
AAQ80916
ID AAQ80916 standard; cDNA; 1320 BP.
XX
XX AAQ80916;
XX
AC 16-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
DT 29-AUG-1995 (first entry)
XX
DE Plasmodium falciparum LSA-1 repeatless gene cDNA.
XX
KM Plasmodium falciparum LSA-1 repeatless gene; recombinant poxvirus;
KM multicomponent multistage malarial vaccines; immunogens;
KM malaria diagnosis; ss.
XX
OS Plasmodium falciparum; (pLSARPLS.14L.1).
XX
PN W09428930-A1.
XX
XX 22-DEC-1994.
PD
XX 10-JUN-1994; 94WO-US006652.
PF
XX 11-JUN-1993; 93US-00075783.
PR 09-JUN-1994; 94US-00257073.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
XX Paolletti E, De Taisne C, Tine JA;
PI
XX WPI, 1995-036113/05.
DR
PT Recombinant poxvirus contg. Plasmodium DNA in non-essential region -
PT useful in vaccines against malaria and for prodn. of Plasmodium
PT immunogens.
XX
XX Claim 3, Fig 11; 183pp; English.
PS
XX
XX AAQ80916 is the *P. falciparum* LSA-1 repeatless gene cDNA sequence. New
CC recombinant poxviruses containing either the SERA, ABRA, Pfhp70, AMA-1,
CC Pf25, Pf26, CSP, PfSP2, LSA-1, LSA-1 repeatless, MSA-1, MSA-1 (N-
CC terminal p83 or C-terminal gp42) genes, or a combination of these in non-
CC essential regions of their genomes are claimed. These poxviruses (pref.
CC with a virulence reducing genomic deletion or disruption) can be used as
CC vaccines against malaria and for the prodn. of Plasmodium immunogens.
CC These viruses provide multicomponent, multistage vaccines due to their
CC expression of sporozite, liver stage, blood stage and sexual stage
CC proteins. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-

CC OCT-2003 to standardise OS field)

XX Sequence 1320 BP, 619 A; 128 C; 227 G; 346 T; 0 U; 0 Other;

5Q Query Match 57.6%; Score 853.2; DB 2; Length 1320;
 Best Local Similarity 87.9%; Pred. No. 1.5e-112;
 Matches 930; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 425 ATTTGACACAGAGACTTGTCTTAAGAAAAGTTGCAGACACACAAAGCGATTGAAAC 484
 DB 263 ATGTAAAAAATGTGCACAAACAAATTCMAAAGCTTTTAAAGAACTTGTGTTTCG 322
 QY 485 AAGAGAGACGTGCTAAAGAAAAGTTGCAGAACCAACAAAGCGATTGAAACAGAGAC 544
 DB 323 AGAATATATCTCTTAAGAAAATTAATTAATTAAGAGGAAATTAATGAAACATTA 382
 QY 545 GTGCTAAAGAAAAGTTGCAGAACCAACAAAGCGATTGAAACAGAGACGTGCTAAAG 604
 DB 383 TAAATGATGATGCGATTAAGAAAATTAATTAATTAAGAGGAAACGAAAACAGACAAAG 442
 QY 605 AAAAGTTGCAGAGCGCAAGAGATTGAAACAAAGGAGCTGATACGAAAAAAAT 664
 DB 443 ATCTTGAAAGAAAAGCGCGCATCTMAAGAAACGAGAAAGCTGATACGAAAAAAAT 502
 QY 665 TGAAGAAAGAAAAGCAATGAGATATATTAGCAGAGATTATATGCTGTTAGAA 724
 DB 503 TAGAAGAAAAGGAAACATGAGATGTTAGCAGAGATTATATGCTGTTAGAA 562
 QY 725 TACAGCTATAGAACTTCCATCAGAAAATGAACGTGATATATATACCATCAATCTT 784
 DB 563 TACAGCTATAGAACTTCCATCAGAAAATGAACGTGATATATATACCATCAATCTT 622
 QY 785 CTTTACCTCAGAGCAACAGAGGAAATAGTAGAGATTCCAAAGAAAATCTTAAATGAAA 844
 DB 623 CTTTACCTCAGAGCAACAGAGGAAATAGTAGAGATTCCAAAGAAAATCTTAAATGAAA 682
 QY 845 AAACAAATAGAGATCTATTACCAACAAATGTTGAAGAGAGGATATATACATTAAGAC 904
 DB 683 AAACAAATAGAGATCTATTACCAACAAATGTTGAAGAGAGGATATATACATTAAGAC 742
 QY 905 ATCTTGAAAGAAAAGAAAGATGTTCAATTAACACAGAACAAAAGAAATTAATCTGCTG 964
 DB 743 ATCTTGAAAGAAAAGAAAGATGTTCAATTAACACAGAACAAAAGAAATTAATCTGCTG 802
 QY 965 ACATTAACAAATCTATCATTTAGAGACAGTAATATTTCTGATGTTAATGATTTCAATTA 1024
 DB 803 ACATTAACAAATCTATCATTTAGAGACAGTAATATTTCTGATGTTAATGATTTCAATTA 862
 QY 1025 GTTAAGATAGAGATTAAGTGCAGAAATAGACGATTCATTAATAGATGAAGAGAG 1084
 DB 863 GTTAAGATAGAGATTAAGTGCAGAAATAGACGATTCATTAATAGATGAAGAGAG 922
 QY 1085 ATGATGAAGACTTAGACGAATTTAGCCTATTGTGCATATAGCAATTTCCAAAGATGAAG 1144
 DB 923 ATGATGAAGACTTAGACGAATTTAGCCTATTGTGCATATAGCAATTTCCAAAGATGAAG 982
 QY 1145 AAAACCTAGGAATTTTAAAGAACTAGAGATTGATAGAGAAAATGAAAAATTTGAGG 1204
 DB 983 AAAACCTAGGAATTTTAAAGAACTAGAGATTGATAGAGAAAATGAAAAATTTGAGG 1042
 QY 1205 ATTTAGATGAAGAAATAGAAAATCATCGAAGAAATTAATCTGAAGAAAATTAAGAAAG 1264
 DB 1043 ATTTAGATGAAGAAATAGAAAATCATCGAAGAAATTAATCTGAAGAAAATTAAGAAAG 1102
 QY 1265 GAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATAAAAAGTTGTATG 1324
 DB 1103 GAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATAAAAAGTTGTATG 1162
 QY 1325 ATGAGCATATTTAAAAATATTAATAATGATAGCAGGTTAATTAAGAAAAAGAAAAATTC 1384
 DB 1163 ATGAGCATATTTAAAAATATTAATAATGATAGCAGGTTAATTAAGAAAAAGAAAAATTC 1222
 QY 1385 TAAATCATTTGTTCAATATATTGACGAGACAAATGAATTTTACAGATGCTGATGAGT 1444

DB 1223 TAAATCATTTGTCATATATTTGACGAGACAAATGAATTTTACAGATCGGATGAGT 1282

QY 1445 TATCTGAAGATATTAATTAATTTTAAAGAACTATTA 1482

DB 1283 TATCTGAAGATATTAATTAATTTTAAAGAACTATTA 1320

RESULT 4

AA005140 ID AA005140 standard; DNA; 1566 BP.

XX AA005140;

AC 25-MAR-2003 (revised)

DT 05-NOV-1990 (first entry)

DE Sequence encoding N-terminal of peptide antigen to malarial sporozite.

XX Malaria; sporozite; vaccine; exoerythrocytic parasites; tetanus toxoid.

XX Plasmodium falciparum.

OS

FT Key Location/Qualifiers

FT CDS 169..1554

FT /*tag= a

PN WO9006130-A.

PD 14-JUN-1990.

XX 30-NOV-1988; 88US-00278234.

PF 30-NOV-1988; 88US-00278234.

XX 30-NOV-1988; 88US-00278234.

PR 12-APR-1989; 89US-00357204.

XX (BIOM-) BIOMEDICAL RES. INST.

PA Hollingdal MR;

PI WPI; 1990-209624/27.

DR P-PSDB; AAR05766.

XX Novel malarial sporozoite antigenic protein - useful as vaccine against

PT sporozoite(s) and exo-erythrocytic parasites.

PS Disclosure; Page ?; -pp; English.

XX Antigen, preferably linked at the C-terminal to a carrier such as tetanus

CC toxoid, may be used as a vaccine against the malarial sporozoite. (Updated

CC on 25-MAR-2003 to correct PI field.)

XX

5Q Sequence 1566 BP; 745 A; 210 C; 330 G; 281 T; 0 U; 0 Other;

Query Match 39.6%; Score 586.4; DB 2; Length 1566;
 Best Local Similarity 91.7%; Pred. No. 1.1e-74;
 Matches 620; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 CAAGAACCAACAAAGCGATCTAGACAGAGACGCTGAAGAAAAGTTGCAAGAACAA 60
 DB 745 CAAGAACCAACAAAGCGATCTAGACAGAGACGCTGAAGAAAAGTTGCAAGAACAA 804
 QY 61 CAAGCGATTTTGAACCAAGATGACTTGTGAAGAAAAGTTTCAAGAGCAAGAGCGAT 120
 DB 805 CAAGCGATTTTGAACCAAGAGACGCTGAAGAAAAGTTTCAAGAGCAAGAGCGAT 864
 QY 121 TTAGAACCAAGAGACCTTGCTAAAGAAAAGTTGCAAGAACCAAGAGCGATAGAACAA 180
 DB 865 TTAGAACCAAGATGACTTGTGAAGAAAAGTTGCAAGAACCAAGAGCGATTTAGAACAA 924
 QY 181 GAGAGACGCTGCTAAAGAAAAGTTGCAAGAACCAAGAGCGATTTAGAACAGAGAGAGCT 240
 DB 925 GATTAAGCGTCTAAAGAAAAGTTGCAAGAACCAAGAGCGATTTAGAACAGATGAGACTT 984

QY 241 CTTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAAAGTAGACTTGTCTAAAGAA 300
DB 985 GCTTAAAGAAAAGTTGCAAGTAACCAAAAGCGATTTAGAACAAAGTAGACTTGTCTAAAGAA 1044
QY 301 AAGTTACAGAGCAGCAAAAGCGATTTAGAACAAAGAGAGAGCGTGTAAAGAAAAGTTGCAAA 360
DB 1045 ACGTTGCAAGAGCCAGCAAAAGCGATTTAGAACAAAGAGAGAGCGTGTAAAGAAAAGTTGCAAA 1104
QY 361 GAAACAAAGCGATTTAGAACAAAGAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAA 420
DB 1105 GAAACAAAGCGATTTAGAACAAAGAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAA 1164
QY 421 AGCGATTTAGAACAAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTA 480
DB 1165 AGAGATTTAGAACAAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTA 1224
QY 481 GAAACAAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAAAGAG 540
DB 1225 GAAACAAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAAAGAG 1284
QY 541 AGAGCGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAAAGAGAGCGTGTCT 600
DB 1285 AGACTGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACAAAGAGAGCGTGTCT 1344
QY 601 AAAGAAAAGTTGCAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGTGAATCAAAA 660
DB 1345 AAAGAAAAGTTGCAAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGTGAATCAAAA 1404
QY 661 AATTAGAAAAGAAA 676
DB 1405 TTGCAAGAACCAAA 1420

RESULT 5

ADO21941
ID ADO21941 standard; DNA; 1371 BP.

ADO21941;

DT 12-AUG-2004 (first entry)

DE LSA-NRC(H) construct DNA derived from Malaria parasite LSA-1.

XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;

KW parasitophorous vacuole; antimalarial; vaccine; immunotherapy;

XX malaria parasite P. falciparum; LSA-NRC(H) construct; ds; gene.

OS Plasmidum falciparum.

XX Synthetic.

XX Key

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

Location/Qualifiers
1..1371
/*tag= a
/product= "LSA-NRC(H) construct protein derived from
Malaria parasite LSA-1"

MO200404167-A2.

27-MAY-2004.

12-NOV-2003; 2003WO-US036011.

12-NOV-2002; 2002US-0425719P.

(REED-) REED ARMY INST RES WALTER.

XX Laner DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;

PI Barbosa A;

XX WPI, 2004-420309/39.

DR P-PSDB; ADO21942.

PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.

PS Claim 16; SEQ ID NO 25; 90pp; English.

XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct DNA of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.

XX Sequence 1371 BP; 539 A; 358 C; 283 G; 191 T; 0 U; 0 Other;

Query Match 34.0%; Score 503.4; DB 12; Length 1371;

Best Local Similarity 63.1%; Pred. No. 6.9e-63; Indels 1; Gaps 1;

Matches 791; Conservative 0; Mismatches 461; Indels 1; Gaps 1;

QY 227 AACAGAGAGAGCGTGTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACCAAGATA 286
DB 79 AACAGAGAAAAGCATGAGAGAAAACATGTCTGAGCCACAACCTCTACGAGAAAGACTAAA 138
QY 287 GACTTGTAAAGAAAAGTTGCAAGAGAGCAAAAGCGATTTAGAACCAAGAGAGCGTCTA 346
DB 139 AACAGAGAAAACCAAAATCTTTGACAAAGAGAGAGCGTCTGAGAACCAAGTTAAA 198
QY 347 AACAGAAAAGTTGCAAGAACCAAAAGCGATTTAGAACCAAGAGAGCGTCTAAAGAA-AA 405
DB 199 AACGATCCAGAGCAAACTTTAAATCCCTCGGCAACTCGAGTTCCGAGAACATC 258
QY 406 TTGCAAGAACCAAAAGCGATTTAGAACCAAGAGAGCGTCTAAAGAAAAGTTGCAAGAA 465
DB 259 TTTCTCAAGAAAACCAAACTGAAACAGAGAGCGCAAACTGTAATCATCAAGAGAC 318
QY 466 CAACAAAGCGATTTAGAACCAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACCAAAAGC 525
DB 319 GAGATGACAAAATAATCATTAAGGCGAGATGAAATCGCCGAGAAAGCTCGAA 378
QY 526 GATTAGAACCAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTTGAA 585
DB 379 GAAAAGAGCTGTGAACGAGCGAGCTGGAACAGAGAGCGCTCGCTAAAGAAAAGCTC 438
QY 586 CAAGAGAGAGCGTGTAAAGAAAAGTTGCAAGAGAGAGAGATTTAGAACCAAAAGAG 645
DB 439 CAGAGAGCGCTCGCTAAAGAAAAGCTCAGAGAGCAAGCGGCGAGCTGGAACGCGAAG 498
QY 646 GCTGATACGAAAAAATTTAGAAAAGAAAAGCAATGAGATATATAGCAGAGAT 705
DB 499 GCTGACAGGAAAAAATCTGGAAGCGAAAAAGGAAACAGGCGAGCTTGCGTGAAGAC 558
QY 706 TTATATAGTGTGTTAGAAATACAGCTATAGAACTTCAATCAAGAAAAGTGAACGTGATAT 765
DB 559 CTGTACGCGCGCTGGAATATCCAGCTATGAACTCCATCCGAAAAGCAAGCGGCTAC 618
QY 766 TATATACCAATGATCTCTTTTACCTCAGAGAACCAAGAGAGGAAATATAGAGATTCAG 825
DB 619 TACATCCACACAGAGAGCGCTGCAAGATATATGCGGAGAACTCCCGGACAGTAG 678
QY 826 GAAATATCTATATAGAAAAAACAATAGAGATCTATTAACAACAAATGTTGAAGAGCA 885
DB 679 GAAATACAGATCATGAAAAAACAACCGGAAAGCATTTCCACCAAGCGTGAAGCGCGC 738
QY 886 AGGATATACATTAAGAGCATCTTGAAGAAAAGAAAGATGTTCAATTAACCAAGAACAA 945
DB 739 CGGACATCCACAAAGCGCACTGGAAGAAAAGAAAGCGCTCCATCAACCAAGAACAG 798
QY 946 AAAGAGATTAATGTGCTGACATCAAAATCATTTAGAGAGAGTAAATATTTCTGAT 1005

```
Db      799  AAGAAGACAAAGCGCTGATATCCAGAACCAACCCCTGGAGACCGTGAAATTAAGCGAC 858
Oy      1006 GTTAAATGATTTTCAAAATAGTAAGTAGGATGAAATTAAGCTGAATATGACGATTCA 1065
Db      859  GTGAACGACTTCAGATCAGCAAGTAGCGAGACGAATCTCCGCTGAATACATGACTCC 918
Oy      1066 TTAATAGTAGAAGAAAGATGATGTAAGACTAGACGAATTTAAGCTTTGTCGAATAT 1125
Db      919  CTGATCGACGAGAGAGAGACGACGAGAGCTTGAGTAATTTCAAAACCAATTTGTCAGTAC 978
Oy      1126 GACAAATTTCCAGATGAGAGAGAAACATAGGAATTTTAAAGAACTAGAGATTTGATAGAG 1185
Db      979  GATTAATTTCAAGACGAGAGAGAAATATCGGCATTTTACAAAGAACTGAGAGACCTCATCGAG 1038
Oy      1186 AAAAAATGAAAATTTAGATGATTTAGTAGAGAGATGAGAAATTCATCAGAGAAATTAATCT 1245
Db      1039 AAAAAAGAAACCTGACGACCTGCGAGAGAGATCGAGAAATTCCTCCAGAGAACTGAGAC 1098
Oy      1246 GAAAGAAAAATTAAGAAAGAAAGAAATTAAGAAAAACAAAGATTAATTTTAAACCA 1305
Db      1099 GAGAGAAAAATCAAAAAAGCAAGAAATCGAAAAACCAAGAGACAACTTCACAAACCA 1158
Oy      1306 AATGATTAAGATTTGTATGATGAGCATATTAATAAATATTAATAATGATAGCAGTTAAT 1365
Db      1159 AAGCAAAATCCCTCTACGACGACGACATTAATAAATTAACAAAACGACAGCAAGTGAC 1218
Oy      1366 AAGGAAAGAGAAAAATTCATTAATATGTTTCAATATTTTGAACGAGACATGAAT 1425
Db      1219 AAGGAAAGAGAAAAATTTATCAAAATCCCTTCACATCTTCGATGCGCATTAACGAAT 1278
Oy      1426 TTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1478
Db      1279 CTGCAAAATGTAGACGACACTGAGCGAGACATCACTAAATTAATCTTCAAGAGCT 1331

RESULT 6
AD021919
ID      AD021919 standard; DNA; 1374 BP.
XX
AC      AD021919;
XX
DT      12-AUG-2004 (first entry)
XX
DE      LSA-NRC(H)Mut construct DNA derived from Malaria parasite LSA-1.
XX
KW      LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KW      parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KW      malaria parasite P. falciparum; LSA-NRC(H)Mut construct; ds; gene;
KW      mutant.
XX
OS      Plasmodium falciparum.
OS      Synthetic.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..1374
FT      /tag= a
FT      /product= "LSA-NRC(H)Mut construct protein derived from
FT      Malaria parasite LSA-1"
XX
XX      MO200404167-A2.
XX
XX      27-MAY-2004.
XX
XX      12-NOV-2003; 2003WO-US036011.
XX
XX      12-NOV-2002; 2002US-0425719P.
XX
XX      (REED-) REED ARMY INST RES WALTER.
XX
XX      Lanat DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W,
XX      Barbosa A;
XX      WPI, 2004-420309/39.
```

```
DR      P-PSDB; AD021920.
XX
PT      Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT      agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT      (LSA-1) epitope.
XX
XX      Claim 17; SEQ ID NO 3; 90pp; English.
XX
CC      The invention relates to a novel recombinant LSA-NRC polypeptide
CC      comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC      protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC      falciparum. The polypeptide of the invention demonstrates antimalarial
CC      activities and may be useful as a vaccine during immunotherapy and as a
CC      diagnostic agent, preferably for diagnosing malaria. The current sequence
CC      is that of the LSA-NRC(H)Mut construct DNA of the invention. The
CC      construct is harmonised for expression in Escherichia coli and comprises
CC      the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC      parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC      within the T5 P. falciparum epitope and a C-terminal His6 tag.
XX
SQ      Sequence 1374 BP; 539 A; 358 C; 285 G; 192 T; 0 U; 0 Other;
XX
Query Match      33.3%; Score 493.6; DB 12; Length 1374;
Best Local Similarity 63.1%; Pred. No. 1.7e-61;
Matches 793; Conservative 0; Mismatches 459; Indels 4; Gaps 2;

227  AACCAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATA 286
Db      79  AACGAGAAAGAGCATGAGAAAGAAACATGCTGACCCCAACTCTACGAGAAAGCTAA 138
Oy      287  GACTGCTAAAGAAAAGTTGCAAGAGCAGCAAGCGATTTAGAACAAAGAGAGCGTGCTA 346
Db      139  AACCAAGAAAACAAATAATCTTTGACAAAGACAAAGAGCTGACATGACAGCAAGTTAA 198
Oy      347  AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTGCTAAAGAA-ANG 405
Db      199  AACGTATCCAGACCAACTTTAAATCCCTCGCCAGCACTCGCGGTTTCCGAGAAATC 258
Oy      406  TTGCAAGAACCAAGAGCATTTAGAACAAAGAGACCTTGCTAAAGAAAAGTTGCAAGAA 465
Db      259  TTTCTCAAGAAACAACTGAAACAAAGAGGCAACATGATTGAACATATCATCAACGAC 318
Oy      466  CAACAAAGCGATTTGAAACAAGAGACGCTGTAAGAAAAGTTGCAAGAACCAAGAAC 525
Db      319  GACGATGACAAAATAATTAATTAAGCCAGATGAAATCCGACAGAAACCTCGAA 378
Oy      526  GATTTAGAACAAAGAGACGCTGTAAGAAAAGTTGCAAGAACCAAGCGATTTAGAA 585
Db      379  GAAAAAGCTGCTGAACAGACATCGGACCTGGAACAGAGCGCTCGCTAAAGAAAAGCTC 438
Oy      586  CAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAGCAGCAAGAGATTTGAACAAAGAA 645
Db      439  CAGGAGCGCTCGCTAAAGAAAAGCTCCAGAGACAAAGCGGACCTGGAACAGCGCAAG 498
Oy      646  GCTGATACGAAAAAATTTAGAAAAGAAAAGAAATGAGATATATTACAGAGAT 705
Db      499  GCTGACAGAAAAAATCTGGAACGCAAAAAGAAACGCGGACGTTCTGCTAGAGAC 558
Oy      706  TTAATGCTGTTTGAATAACAGCTATAGAACTTCATCAGAAAATGAACGCTGATAT 765
Db      559  CTGTAAGCGCGCTGGAATATCCAGTATTCGAATCCCATCCGAAACGAAAGCGGCTAC 618
Oy      766  TATATACCAATCAATCTTTTACCTCAGACCAACAGAGGAATATAGATTCGAAG 825
Db      619  TACATCCCAACACAGAGAGCCTGCAAGATTAATCGGGGAATCCCGCAGAGTAAG 678
Oy      826  GAAATATCTATATAGAAAAAACAATAGAGATCTATTACAACAATTTGAGAGACGA 885
Db      679  GAAATCAGCATATGAAAAAACAACGCGAGAAAGCATTTACCAACAGCTGAGAAAGCGCC 738
Oy      886  AAGGATATATCAATAAGACATCTTGAAGAAAAGAAAGATGTTCAATAAACAAGAACAA 945
Db      739  CGGACATCCCAAGAGGCACTCGAAGAAAAAGAAAGACGGCTCATCAACCAAGAACAG 798
```

```
QY 946 AAAGAAGTAATCTGCTACATACAAATCATATAGAGACGTAATTTTCAT 1005
DB 799 AAAGAAGCAAAAGCGCTGATATCCAGAACACCCCTGGACCGTAACATTAGCGAC 858
QY 1006 GTTAATGATTTTCAATTAAGTAATAGAGATGAATTAAGTCTGAATATGACGATTTCA 1065
DB 859 GTGAACGACTTCCAGATACGAAAGTACGAGAACGAAATCTCCGCTGAATACGATGACTCC 918
QY 1066 TTAATAGATGAAGAAAGATGATGAGACTTACGCAATTTAAAGCTATTGTGCAATAT 1125
DB 919 CTGATCGACGAAGAAAGACGACGAGATCTGATGATTAATCAACCAATGTCTCAGTAC 978
QY 1126 GACAAATTTCCAAAGTAAGAAACATGAGAAATTTATAAGAACTAAGAAATTTGATGAG 1185
DB 979 GATTAACCTTTCAGACGAAAGAAATATTCGGCTTTTACAAAGAACTCGAAGACCTCATCGAG 1038
QY 1186 AAAATGAAATTTAGATGATTTAGATGAAGAAATAGAAATATCATCAGAAATTAATCT 1245
DB 1039 AAAAAGCAAAACCTGACGACCTGACGAAAGCATCGAAATATCTCCGAAGAACTGAC 1098
QY 1246 GAAGAAAAAATTAATAAGAAATATGAAAAAACAAGATTAATTTTAAACCA 1305
DB 1099 GAAGAAAAAATCAAAAAAGCAAGAAATACGAAAAACCAAGACCAACTTCAACCA 1158
QY 1306 AATGATAAAGATTTGTATGATGACATATTAATAAATTAATAAATGATTAAGAGTTAAT 1365
DB 1159 AACGACAAATCCCTCTACGACGACACATTAATAAATTAATAAATGACAAAGTGAAC 1218
QY 1366 AAGGAAAAAGAAAAATTCATTAATTCATTTTCAATATTTGACGAGCAATGAATTT 1425
DB 1219 AAGGAAAAAGAAAAATTTATCAATTCCTCTTCCACATCTTCGATGGGGAATGAATTT 1278
QY 1426 TTACAGATCGTGATGA---GTTATCTGAAGATATTAATAAATTTTATGAATCT 1478
DB 1279 CTGCAAAATTTGACGAAACGTTGAGCAAGACATCACTAATATCTTCATGAACT 1334

RESULT 7
AAQ28117 standard; DNA; 988 BP.
ID AAQ28117 standard; DNA; 988 BP.
AC AAQ28117;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum LSA gene 5' region.
XX
KM Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen; ss.
XX
OS Plasmodium falciparum.
XX
XX Key location/Qualifiers
FH 5'UTR 1..32
FT /tag= a
FT /tag= a
FT /tag= d
FT /note= "5' part of LSA gene"
FT repeat_region 492..988
FT /tag= b
FT repeat_unit 492..542
FT /tag= c
XX
XX WO9213884-A1.
XX
XX 20-AUG-1992.
XX
XX 05-FEB-1992; 92WO-FR000104.
XX
XX 05-FEB-1991; 91FR-00001286.
XX
```

```
PA (INSP ) INST PASTEUR.
XX
XX PI
XX Guertimarchand C, Drullhe P;
XX
XX WPI; 1992-299985/36.
DR P-PSDB; AAR26943.
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 11; Fig 6; 81pp; French.
PS
XX
XX This nucleotide sequence is the 5' part of the P.falciparum liver-stage
CC specific antigen (LSA) gene. It codes for a polypeptide sequence which
CC carries a T cell epitope characteristic of a protein produced in
CC hepatocytes infected with P.falciparum. The polypeptide can be used in
CC the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
XX Sequence 988 BP; 465 A; 120 C; 184 G; 219 T; 0 U; 0 Other;

Query Match 31.9%; Score 472.8; DB 2; Length 988;
Best Local Similarity 95.7%; Pred. No. 1.5e-58;
Matches 486; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 77 AAGTACCTTGTAAAGAAAGTTTACAGAGCAGCAAGCGATTTAGAACAGAGAGAC 136
DB 481 AAGAAAAAGCAAGCTTAAGAAAGTTTACAGAGGCAACAAAGCATTCAGAACAGAGAGAC 540
QY 137 TTGCTAAAGAAAGTTGCAAGAACCAACAAAGCATTTAGAACAGAGAGAGAGAG 196
DB 541 GTGCTAAAGAAAGTTGCAAGAACCAACAAAGCATTTAGAACAGAGAGAGAGAGAG 600
QY 197 AAAAGTTGCAAGAACCAACAAAGCATTTAGAACAGAGAGAGAGAGAGAGAGAGAG 256
DB 601 AAAAGTTGCAAGAACCAACAAAGCATTTAGAACAGAGAGAGAGAGAGAGAGAGAG 660
QY 257 AAGAACCAACAAAGCGATTTGAAACAAAGTACGCTGCTAAAGAAAGTTTCAAGAGAGC 316
DB 661 AAGAACCAACAAAGCGATTTGAAACAAAGTACGCTGCTAAAGAAAGTTTCAAGAGAGC 720
QY 317 AAGCGATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 376
DB 721 AAGCGATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 377 TAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
DB 781 TAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 437 AGAGACTTGTAAAGAAAGTTGCAAGAACCAACAAAGCGATTTAGAACAGAGAGAGAG 496
DB 841 AGAGACTTGTAAAGAAAGTTGCAAGAACCAACAAAGCGATTTAGAACAGAGAGAGAG 900
QY 497 CTAAAGAAAGTTGCAAGAACCAACAAAGCGATTTAGAACAGAGAGAGAGAGAGAG 556
DB 901 CTAAAGAAAGTTGCAAGAACCAACAAAGCGATTTAGAACAGAGAGAGAGAGAGAG 960
QY 557 AGTTGCAAGAACCAACAAAGCGATTTAGA 584
DB 961 GGTGCAAGAACCAACAAAGCGATTTAGA 988

RESULT 8
ADP85917 standard; DNA; 1300 BP.
ID ADP85917 standard; DNA; 1300 BP.
AC ADP85917;
XX
XX 26-AUG-2004 (first entry)
XX
XX Synthetic construct #1.
DB
XX Nanopore data analysis ; polymer ; polynucleotide;
XX
```



```
OY 1351 GATAAGCAGTTTATAGAGAAAGAAATTCATPAAA 1389
DB 1261 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1299

RESULT 9
AA05868
XX AA05868 standard; DNA; 3399 BP.
XX
AC AA05868;
XX
DT 27-AUG-2003 (revised)
DT 14-AUG-1996 (first entry)
XX
DE Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.
XX
KM Chicken leucocytozoan; immunogen; recombinant vaccine; protection;
XX Immunisation; vaccination; ss.
XX
OS Leucocytozoan BP.
XX
FH Key Location/Qualifiers
FH CDS 1..3399
FT /*tag= a
FT mlec_feature 1150..3218
FT /note= "fragment referred to in the claims, for use as
FT insert in a recombinant vaccine against chicken
FT leucocytozoan disease"
XX
PN JP07284392-A.
XX
PD 31-OCT-1995.
XX
PF 19-APR-1994; 94JP-00080643.
XX
PR 19-APR-1994; 94JP-00080643.
XX
PA (DOBU-) DOBUTSUYO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.
PA (KITA) KITASATO KENKYUSHO SH.
XX
DR WPI: 1996-006311/01.
DR P-PSDB: AAR97866.
XX
PT Chicken leucocytozoan immunogenic protein - used in a recombinant vaccine
PT against chicken leucocytozoan disease.
XX
PS Claim 6; Page 6-9; 35pp; Japanese.
XX
CC AA05868 encodes a chicken leucocytozoan immunogenic protein, this DNA or
CC a fragment of it can be used in a recombinant vaccine to immunise against
CC chicken leucocytozoan disease. The DNA is used in a vector and
CC operatively linked to an expression regulatory sequence as in standard
CC practice. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 3399 BP; 1577 A; 508 C; 798 G; 516 T; 0 U; 0 Other;

Query Match 15.2%; Score 225.8; DB 2; Length 3399;
Best Local Similarity 48.7%; Pred. No. 1.6e-23;
Matches 713; Conservative 0; Mismatches 737; Indels 15; Gaps 3;

OY 2 AAGAACAAACAAGCGATTAGAACAGAGACGTGCTPAAAGAAAAGTTGCAAGAACAC 61
DB 1685 AAGTAAACATGAGAAATAGAAAAAGAGACATGAAGAAATATACATGAGAAAGAA 1744
OY 62 AAAGCATTTAGAACAAATAGACTTGTAAAGAAAGTTTCAAGAGACAGAAAGCGATT 121
DB 1745 AAGAAAGAGTAAACATAGAAATAGAAAAAGAGACATGAGAAATATACATGAG 1804
OY 122 TAGAACAAAGAGACTTCTAAAGAAAAGTTGCAAGAACAAAGCAGCTAGAACAG 181
DB 1805 AAGAAAAAGAAAGTAAACATGAGAG-----AAGAAAAAGAAAGACATGAGAG 1858
```

```
OY 182 AGAGACGTGCTAAAGAAAAGTTGCCAAGAACAAACAAAGCGATTAGAACAGAGACGTG 241
DB 1859 TAATACATGAGAAAGAAAAGAAAGTAACACATGAGAAATAGAAAAAGAGCATG 1918
OY 242 CTAAAGAAAAGTTGCCAAGAACAAACAAAGCGATTAGAACAGATGACTTGTAAAGAAA 301
DB 1919 AAGAAAGTAAATCATGAGAAAGAAAAGAAAGAGTAAACATGAGAAATAGAAAAAGAG 1978
OY 302 AGTTACAAAGACGCAAAAGCGATTAGAACAGAGACGCTTAAGAAAAAGTTGCAAG 361
DB 1979 AGCATGAAAGAGTAAATCATGAGAGAAAAAGAAAGTAAACATGAGAAATAGAAA 2038
OY 362 AACAAACAAGCGATTAGAACAGAGACGCTGCTAAAGAAAAGTTGCCAAGAACAAACA 421
DB 2039 AAGAA-----GACATGAAAGAGTAAATCATGAGAGAAAAAGAAAGAGTAAACATG 2092
OY 422 GCGATTAGAACAAAGAGACGTTGCTAAAGAAAAGTTGCCAAGAACAAACAAAGCGATTAG 481
DB 2093 AAGAAATAGAAAAGAAAGAGACATGAGAAAGTAAATCATGAGAAAGAAAAGAAAGTAA 2152
OY 482 AACAAAGAGACGCTGCTAAAGAAAAGTTGCCAAGAACAAACAAAGCGATTAGAACAAAGA 541
DB 2153 CACATGAAAGAAAAGAAAAGAAAGACATGAGAAAGTAAATCATGAGAAAGAAAAGAG 2212
OY 542 GACGTGCTAAAGAAAAGTTGCCAAGAACAAACAAAGCGATTAGAACAGAGACGCTGCTA 601
DB 2213 AAGTAAACATGAGAAAGAAAGAAAGAAAGTAAACATGAGAAAGAAAGAAAGTAA 2272
OY 602 AAGAAAAGTTGCCAAGACAGCAAAAGATTTAGAACAAAGAAAGCGTGAATACGAAAAAA 661
DB 2273 TACATGAAAGAAAGAAAAGAAAGTAAATCATGAGAAAGAAAAGAAAGTAAACATG 2332
OY 662 ATTTAGAAAGAAAAAGAAACATGAGATTAATTTAGCAGAGATTTATATGTCGTTAG 721
DB 2333 AAGAAAGAAAAGAAAGAAAGTAAACATGAGAAAGAAAAGAAAGTAAACATGAGAGAG 2392
OY 722 AATATCAGCGTATAGAACTTCATCAGAAAATGAAACGTGATATATATATACCATCAAT 781
DB 2293 AAAAAAGAAAGTAAACATGAGAAAGAAAGAAAGTAAACATGAGAAAGAAAGTAA 2452
OY 782 CTCTTTACCTCAGAGCAACAGAGGAATAGAGATTCAGAGAAATATCTTAATAG 841
DB 2453 AAGTAAACATGAGAAAGAAAGAAAGAAAGTAAACATGAGAAAGAAAGAAAGTAA 2512
OY 842 AAAAAACAATAGAAATCTATTCAACAAATGTTGAAGCAAGAGGATTAATTAAG 901
DB 2513 CACATGAAAGAAAGAAAAGTAAACATGAGAAAGAAAGTAAACATGAGAGAG 2572
OY 902 GACATCTTGAAGAAAGAAAGATGCTCAATTAACAGAACAAAGAAAGATTAATCTG 961
DB 2573 AAGAAAAGTAAACATGAGAAAGAAAGAAAGAAAGTAAACATGAGAGAG 2632
OY 962 CTGACATACAAAATCATATACATTTAGAGACAGTAATATTTCTGATGTTAATGATTTTCAA 1021
DB 2633 AAGAAAAGTAAACATGAGAAAGAAAGAAAGTAAACATGAGAGAGAAAGAAAGTAA 2692
OY 1022 TAACTAAGTATAGC---ATGAAATTAAGTCTGAATTAATGACATTCATTAATAGTAGAG 1078
DB 2693 TACATGAAAGAAAGAAAAGAAAGAAAGTATGAGAAAGAAAGAAAGAAAGAG 2752
OY 1079 AAGAAAGTGAATGAGAGACTTGAAGCAATTTAAGCCTAATTTGCAATATGACAAATTTCCAG 1138
DB 2753 AAGAAAGAAAGAAAGAAAGATGAGAGAGAGAAAGAAAGAAAGAAAGATGAGAGAG 2812
OY 1139 ATGAAGAAAACATAGAAATTTAATAAGAACTAGAGAGTTTGTATGAGAAAAATGAATTT 1198
DB 2813 AAGAAAGAAAGAAAGAAATGAGAGAGAAAGAAAGAAAGAAATTAAGAAAGAGAGAG 2872
OY 1199 TAGATGATTTAGATGAAGAAATAGAAAATCATCAGAGAAATTTCTGAAGAAAAATTA 1258
DB 2873 AAGAAAAAGAAAGCAATGAGAAAGAAAGTAAACATGAGAAAGAAAGAAAGTAAACAC 2932
OY 1259 AAAAAAGAAAGAAATATGAGAAAAAGAAAGATATATTTTAAACCAATGATAAAGTT 1318
```

```

Db      2933  ATGAAGAAGAAAAAGTAACACATGAAGAGAAAGAAATGTAACATATGAAGAAG 2992
QY      1319  TGTATGATGAGCATTTTAAAAATATAAAAATGATTAAGAGTTAATAGAAAAAGAA 1378
Db      2993  AAGAAAAAGTAAACACATGAAGAAAGAAAGTAACATATGAAGAGAAAAAGTAA 3052
QY      1379  AATTGATTAATCATTTGTTTCATATATTTGACGAGACATGAAATTTTACAGATCGTGG 1438
Db      3053  CACATGAAGAGAGAAATGTAACATATGAAGAGAGAGAGAAAAAGTAAACACATGAAG 3112
QY      1439  ATGAGTTATCTGAAGATTAAGTAA 1463
Db      3113  AAGAGAGAGAAAAAGTAAAGAA 3137

RESULT 10
AAN80899 standard; DNA; 210 BP.
XX
AC      AAN80899;
XX
DT      14-MAY-2003 (revised)
DT      25-MAR-2003 (revised)
DT      19-SEP-1990 (first entry)
XX
DE      Sequence encoding an epitope characteristic of a protein produced in
DE      liver cells infected with Plasmodium falciparum.
XX
KW      Plasmodium falciparum; vaccine; malaria diagnosis;
KW      malaria antibody detection.
XX
OS      Plasmodium falciparum.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..204
FT      /tag= a
XX
PN      MO8805785-A.
XX
PD      11-AUG-1988.
XX
PF      09-FEB-1988; 88WO-FR000074.
XX
PR      09-FEB-1987; 87PR-00001543.
XX
PA      (INSP ) INST PASTEUR.
PA      (MARC/) MARCHAND C.
PA      (CNRS ) CENT NAT RECH SCI.
XX
PI      Druilhe P, Pujjalome O, Marchand C, Langeley G;
XX      WPI; 1988-235148/33.
XX      P-PSDB; AAP81355.
XX
PT      New polypeptide(s) contg. plasmodium falciparum epitope - useful for
PT      malaria vaccination or diagnosis.
XX
PS      Example 1; Fig 1; 37pp; French.
XX
CC      This sequence is contained in clone DG307 of a genomic DNA library of P.
CC      falciparum. AAN80899 is composed entirely of a repeating motif of 51 base
CC      pairs. The type of repetition coded by DG307 seems to be conserved which
CC      implies that it may be present in all P. falciparum strains. The AA
CC      repeat unit is the basis of the peptides claimed in the patent, which are
CC      useful for the prodn. of of malaria vaccines and as immunoassay reagents
CC      for diagnosis of malaria by detection of antibodies. (Updated on 25-MAR-
CC      2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
CC      (Updated on 14-MAY-2003 to correct PS field.)
XX
SQ      Sequence 210 BP; 97 A; 33 C; 47 G; 33 T; 0 U; 0 Other;
Query Match      12.4%; Score 183.2; DB 1; Length 210;

```

```

Beat Local Similarity 95.9%; Pred. No. 2.2e-17;
Matches 188; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY      83  GACTTGCTAAAGAAAGTTACAGAGCAGCAAGCGATTTAGAACAAAGAGACTTGGCTA 142
Db      8  GACTTGCTAAAGAAAGTTACAGAGCAGCAAGCGATTTAGAACAAAGAGACTTGGCTA 67
QY      143  AAGAAAGTTCCAGAACAAAGAAAGCGATCTAGAACAAAGAGAGCGTCTAAAGAAAGT 202
Db      68  AAGAAAGTTCCAGAACAAAGAAAGCGATCTAGAACAAAGAGAGACTTAAAGAAAGT 127
QY      203  TGCAGAAACAAAGCGATTTAGAACAGAGAGCGTCTAAAGAAAGTTCGCAAGAAC 262
Db      128  TACAGGGGCAACAAAGCGATCTAGAACAAAGAGAGCTTCTAAAGAAAGTTGCAAGAAC 187
QY      263  AACAAAGCGATTTAGA 278
Db      188  AACAAAGCGATTTAGA 203

RESULT 11
ADO62833 standard; RNA; 1000 BP.
XX
AC      ADO62833;
XX
DT      07-OCT-2004 (first entry)
XX
DE      Homopoly-A contaminant for RNaseH activity assay.
XX
KW      ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;
KW      fluorophore; fluorescence quencher; fluorescent signal; HIV;
KW      reverse transcriptase.
XX
OS      Synthetic.
XX
PN      WO2004059012-A1.
XX
PD      15-JUL-2004.
XX
PF      22-DEC-2003; 2003WO-US040879.
XX
PR      23-DEC-2002; 2002US-0436125P.
XX
PA      (AMHP ) WYETH.
XX
PI      Olson MW, O'Connell JF;
XX      WPI; 2004-543471/52.
XX
PT      Detecting a nuclease-mediated cleavage of a target nucleic acid, useful
PT      for detecting and monitoring RNase H activity, comprises hybridizing a
PT      target nucleic acid to a fluorescently labeled oligonucleotide probe.
XX
PS      Example 1; SEQ ID NO 4; 61pp; English.
XX
CC      The invention relates to a method of detecting a nuclease-mediated
CC      cleavage of a target nucleic acid or measuring a RNaseH activity of an
CC      agent by hybridizing a target nucleic acid to a fluorescently labelled
CC      oligonucleotide probe complementary to the target nucleic acid and
CC      containing a fluorophore at one terminus and a quenching group at the
CC      other terminus and contacting the probe-target hybrid with an agent
CC      having nuclease activity. When the oligonucleotide probe is unhybridised
CC      to the target nucleic acid, the probe adopts a conformation that places
CC      the fluorophore and quencher in such proximity that the quencher quenches
CC      the fluorescent signal of the fluorophore. Formation of the probe-target
CC      hybrid causes sufficient separation of the fluorophore and quencher to
CC      reduce quenching of the fluorescent signal of the fluorophore followed by
CC      contacting the probe-target hybrid with an agent having nuclease activity
CC      to selectively cleave the target nucleic acid and thus release the intact
CC      probe. Detection of the release of the probe is by measuring a decrease
CC      in the fluorescent signal of the fluorophore as compared to the signal of
CC      the probe-target hybrid. The methods are useful for detecting and

```

CC monitoring RNaseH activity in real time. The methods are also useful for
 CC monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
 CC the invention, the activity of the RNaseH from E. coli or HIV reverse
 CC transcriptase is measure by the method of the invention. In order to test
 CC the specificity of the RNaseH cleavage of the target-probe complex, a DNA
 CC or RNA contaminant is added to the assay reaction. This sequence
 CC corresponds to the single stranded homopolymeric polyA RNA contaminant
 CC for testing this activity.

XX Sequence 1000 BP; 1000 A; 0 C; 0 G; 0 T; 0 U; 0 Other;

Query Match 12.2%; Score 180.2; DB 12; Length 1000;

Best Local Similarity 48.8%; Pred. No. 5.3e-17;

Matches 485; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

```

Qy 397 AAGAGAAAGTTGCAAGACAAAGCGATTAGACAGAGAGACTGTCAAGAAAG 456
    |||||
Db 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60

Qy 457 TTGCAGACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAA 516
    |||||
Db 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120

Qy 517 CAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAAGTTGCAAGAAAGCAAGC 576
    |||||
Db 121 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180

Qy 577 GATTAGACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAGCAAGAGATTAGAA 636
    |||||
Db 181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 240

Qy 637 CAAGAGAGCGTCTAGCAAGAAAGTTAGAAAGAAAGAACTGAGATATTTA 696
    |||||
Db 241 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300

Qy 697 GCAGAGATTATATGCTTTAGAAATACACGCTATAGAACTCCATCAGAAATGAA 756
    |||||
Db 301 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 360

Qy 757 CGTGATTTATATACCATCAATCTTCTTACCTCAGACAAAGAGGATATAGAA 816
    |||||
Db 361 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420

Qy 817 GATTCCAGAGAAATATCTATATAGAAAGAAAGAAATGAGATCTTACAAATGTT 876
    |||||
Db 421 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 480

Qy 877 GAAGAGCAAGGATATATCATTAAGGACATCTTGAAAGAAAGAAAGATGTTCAATAAA 936
    |||||
Db 481 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 540

Qy 937 CCAGACAAAGAGAAATATCTGTCATACAAATCATACATTAGACAGATTAAT 996
    |||||
Db 541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600

Qy 997 ATTTCTGATGTTATGATTTTCAATAGTAAGATGAGATGAATATAGTCTGAATAT 1056
    |||||
Db 601 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 660

Qy 1057 GACGATTTATATATGATGAAGAGATGATGAGCTTATGAGCAATTTTAAAGCTATTT 1116
    |||||
Db 661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 720

Qy 1117 GTGCATATGACAAATTTCCAGATGAGAAAGATGAGAAATTTTAAAGACTAGAGAT 1176
    |||||
Db 721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 780

Qy 1177 TTGATAGAGAAAGATGAGAAATTTAGATGATTTAGATGAGAAATGAGAAATTCATCAGAA 1236
    |||||
Db 781 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 840

Qy 1237 GAATTTATCTGAGAAAGAAATTAAGAAAGAAATGAGAAAGAAATGAGATTAATAT 1296
    |||||

```

```

Db 841 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900

Qy 1297 TTTAAACCAATATATATTTGATGACATTTTAAATATATATATGATTAAG 1356
    |||||
Db 901 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 960

Qy 1357 CAGTTAATGAGAAAGGAAATTTCAATTAATA 1389
    |||||
Db 961 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 993

```

RESULT 12

AD062832/c

ID AD062832 standard; RNA; 1000 BP.

XX AD062832;

XX 07-OCT-2004 (first entry)

XX Homopoly-U contaminant for RNaseH activity assay.

XX ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;

KM fluorophore; fluorescence quencher; fluorescent signal; HIV;

KM reverse transcriptase.

OS Synthetic.

XX WO2004059012-A1.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003NO-US040879.

XX 23-DEC-2002; 2002US-0436125P.

XX (AMHP) WYETH.

PI Olson MM, O'Connell JF;

XX WPI: 2004-543471/52.

XX Detecting a nuclease-mediated cleavage of a target nucleic acid, useful

XX for detecting and monitoring RNase H activity, comprises hybridizing a

XX target nucleic acid to a fluorescently labeled oligonucleotide probe.

XX Example 1; SEQ ID NO 3; 61pp; English.

CC The invention relates to a method of detecting a nuclease-mediated
 CC cleavage of a target nucleic acid or measuring a RNaseH activity of an
 CC agent by hybridizing a target nucleic acid to a fluorescently labelled
 CC oligonucleotide probe complementary to the target nucleic acid and
 CC containing a fluorophore at one terminus and a quenching group at the
 CC other terminus and contacting the probe-target hybrid with an agent
 CC having nuclease activity. When the oligonucleotide probe is unhybridised
 CC to the target nucleic acid, the probe adopts a conformation that places
 CC the fluorophore and quencher in such proximity that the quencher quenches
 CC the fluorescent signal of the fluorophore. Formation of the probe-target
 CC hybrid causes sufficient separation of the fluorophore and quencher to
 CC hybrid quenching of the fluorescent signal of the fluorophore followed by
 CC contacting the probe-target hybrid with an agent having nuclease activity
 CC to selectively cleave the target nucleic acid and thus release the intact
 CC probe. Detection of the release of the probe is by measuring a decrease
 CC in the fluorescent signal of the fluorophore as compared to the signal of
 CC the probe-target hybrid. The methods are useful for detecting and
 CC monitoring RNaseH activity in real time. The methods are also useful for
 CC monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
 CC the invention, the activity of the RNaseH from E. coli or HIV reverse
 CC transcriptase is measure by the method of the invention. In order to test
 CC the specificity of the RNaseH cleavage of the target-probe complex, a DNA
 CC or RNA contaminant is added to the assay reaction. This sequence
 CC corresponds to the single stranded homopolymeric polyU RNA contaminant
 CC for testing this activity.

Sequence 1000 BP; 0 A; 0 C; 0 G; 0 T; 1000 U; 0 Other;

Query Match 12.2%; Score 180.2; DB 12; Length 1000;
 Best Local Similarity 48.8%; Pred. No. 5,3e-17;
 Matches 485; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

```

QY 397 AAAAAAAAAAGTTCCAGAACCAACAAACGCGATTAGAACAGAGAGACTTGCTTAAGAAAAG 456
DB 1000 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 941
QY 457 TTGCAAGAACCAACGCGATTAGAACAGAGAGCGCTTAAGAAAAGTTGCAAGAA 516
DB 940 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 881
QY 517 CAACAAGCGATTAGAACAGAGAGCGCTTAAGAAAAGTTGCAAGAACCAAGC 576
DB 880 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821
QY 577 GATTAGAACAGAGAGCGCTTAAGAAAAGTTGCAAGAGCAGCAAGAGATTAGAA 636
DB 820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
QY 637 CAAGAGCGCTGATACGAAAAAAAAATTAGAAAGAAAAGCAATGAGATATATTA 696
DB 760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
QY 697 CGAGAGGATTTATATGCGCTTGAATAACAGCTATAGAACTTCATCAGAAAATGAA 756
DB 700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641
QY 757 CGTGATATTATATACACATCAATCTCTTACCTCAGAGCAACAGAGGATATAGAA 816
DB 640 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 581
QY 817 GATTCAGAGAAATATCTATATAGAAAAACAATAGAAATCTATTAACAATGTT 876
DB 580 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 521
QY 877 GAAAGCAGAGGATATATATAGAGACATCTTGAAGAAAGAAAGATGTTCAATAAA 936
DB 520 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 461
QY 937 CCAGAACCAAAAGAGATTAATCTGCTGACATACAAATCATCATTAAGACAGTAAT 996
DB 460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
QY 997 ATTTCTGATGTTAATGATTTTCAAAATAGTATGAGATGAATTAAGTCTGATAT 1056
DB 400 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341
QY 1057 GACGATCTTATATAGATGAGAGAGATGATGAAGACTTAGACGAATTTAAGCTATT 1116
DB 340 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 281
QY 1117 GTGCAATATGACAAATTTCCAGATGAGAAACATAGGAATTTATTAAGACTAGAGAT 1176
DB 280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221
QY 1177 TTGATAGAGAAAATGAAAATTTAGATGATTAGATGAAGAAATAGAAAATCATCAGA 1236
DB 220 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 161
QY 1237 GAATTTCTGAGAGAAAATTAAGAAAAGAGAAATATGAAAACCAAGATTAATAT 1296
DB 160 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101
QY 1297 TTTAAACCAATGATTAAGTTTGTATGATGACATATTAAAAATATATAATGATAAG 1356
DB 100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 41
QY 1357 CAGGTTAATAGGAAAAGAAAATTCATAAA 1389
DB 40 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

```

RESULT 13

AA087587
 ID AA087587 standard; DNA, 1686 BP.

XX AA087587;

XX AC 27-AUG-2003 (revised)
 XX DT 19-DEC-1995 (first entry)

XX DE DNA encoding Leucocytozoan protozoa structural protein epitope.

XX KW leucocytozoan protozoa; structural protein; epitope; vaccine; fowl;
 XX leucocytozoanosis; treatment; ss.

XX OS Leucocytozoan.

XX PN JP07089995-A.

XX PD 04-APR-1995.

XX PF 10-SEP-1993; 93JP-00226078.

XX PR 10-SEP-1993; 93JP-00226078.

XX PA (DOBU-) DOBUTSUYO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.
 XX (NISS-) NISSEIKEN KK.

XX DR WPI, 1995-167252/22.

XX DR P-PsDB; AAR70491.

XX PT Immune inducing polypeptide against Leucocytozoan protozoa - useful in
 XX production of vaccines for treatment of leucocytozoanosis in fowl.

XX PS Claim 1; Page 12-14; 20pp; Japanese.

CC AA087587-69 encode polypeptides having a whole or partial epitope of a
 CC structural protein of Leucocytozoan protozoa (see AAR70491-93). The
 CC polypeptides and DNA encoding them are useful in the production of
 CC vaccines for the treatment of Leucocytozoanosis of fowl. (Updated on 27-
 CC AUG-2003 to correct OS field.)

XX Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 U; 0 Other;

Query Match 12.1%; Score 179.4; DB 2; Length 1686;
 Best Local Similarity 46.9%; Pred. No. 6,6e-17;
 Matches 595; Conservative 0; Mismatches 671; Indels 3; Gaps 1;

```

QY 2 AAGAACCAAAAGCGATCTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAC 61
DB 101 AAGAGAAAATGAGAAAATGAAAATGAAAAGAAACAAGAGAGAGAAACAAGAACAAAG 160
QY 62 AAAGCGATTAGAACAGATAGACTTGCTTAAGAAAAGTTACAGAGCAGCAAGCGATT 121
DB 161 AACAGAACCAAGAAATCGTAGAAGAAACAAGAACAGATGAAGAAACAAGAGAGAAAG 220
QY 122 TAGAACAGAGAGACTTGCTTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAG 181
DB 221 ATGAAGAGAGAGAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 280
QY 182 AGAGCGTGTAAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACCAAGAGAGAGCGTG 241
DB 281 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
QY 242 CTAAAGAAAAGTTGCAAGAACCAAGCGATTTAGAACCAAGATGACTTGCTTAAGAAA 301
DB 341 ATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
QY 302 AGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
DB 401 AAAATGAAGATGAACAAATGAGATGAACAAATGAGAGAGAGAGAGAGAGAGAGAGAG 460

```

```
QY 362 AACACAAAGCGATTTAGAACAGAGAGAGCTGTAAGAAAGTTGCAAGAACAA 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 461 AAGAACAAACAAAGATGAAAGAGAAACAAAGTAAGAAACAAAGTAAGAGAG 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 422 GCGATTTAGAAAGAGAGACTTGTCTAAGAAAGTTGCAAGAACAAAGCGATTTAG 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 521 AAGAACAAAGAGAGAGAGAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 482 AACAGAGAGAGCTGTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAG 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 581 ATGAAGATGAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 542 GAGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGAGAG 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 641 AAAAAAGAAAGAGAGAGAGATGTAAGAAAGAGAGAGAGAGAGAGAGAGAGAG 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 602 AAGAAAGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 701 AAAAAAGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 662 ATTTAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 758 CAAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 817
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 722 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 818 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 877
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 782 CTTCTTAACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 878 CAAATGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 937
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 842 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 938 TAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 997
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 902 GAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 961
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 998 ATGAATCAAAATCCGAATTAATTCATCAATTAATTTGAAGAGAGAGAGAGAG 1057
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 962 CTGACATACAAATCATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1058 CATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1022 TAACTAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1118 TTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1082 AAGATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1178 TAGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1142 AAGAAACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1238 CACCAAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1202 ATGATTTAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1298 AAGAAACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1262 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1358 TAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:232.
XX
XX Plasmodium falciparum; chromosome 2, human malaria parasite; vaccine;
KW antimalarial; malaria; protozoa; infection; insecticide; ds.
XX
XX Plasmodium falciparum.
XX
XX MO200025728-A2.
XX
XX 11-MAY-2000.
XX
XX
XX 05-NOV-1999; 99MO-US026796.
XX
XX 05-NOV-1998; 98US-0107131P.
XX
XX (HOFF/) HOFFMAN S.
XX (CARU/) CARUCCI D.
XX (GARD/) GARDNER M.
XX (VENT/) VENTER J C.
XX
XX Hoffman S, Carucci D, Gardner M, Venter JC;
XX WPI; 2000-365347/31.
XX
XX
XX The present invention describes proteins and their fragments (I) encoded
XX by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
XX Also described are: (I) nucleotide sequences (I) encoding (I); and (2)
XX vaccines against P. falciparum infection comprising (I) or (II). (I) and
XX (II) are useful for the development of vaccines against P. falciparum
XX infection. (I) and polyclonal antisera or a monoclonal antibody raised to
XX immunogens comprising the sequences of (I), are useful in the detection
XX of infection with P. falciparum. Furthermore, (I) (especially when they
XX are rRNAs or secreted or membrane proteins) can aid the identification
XX of drugs to treat or prevent P. falciparum infection, or they can be used
XX to identify drug resistance in P. falciparum. Sequencing of the
XX Plasmodium chromosome 2 and the subsequent identification of proteins
XX encoded by it will help to expand our understanding of parasite biology,
XX a process hampered by the complexity of the parasite life cycle, and
XX provide new targets for vaccine and drug development. Parasite resistance
XX to drugs and mosquito resistance to insecticides have led to a resurgence
XX of malaria in many parts of the world, and there is a pressing need for
XX vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
XX represent nucleotide and protein sequences given in the present
XX invention, but which are not specifically mentioned within the
XX specification
XX
XX
XX Sequence 3579 BP; 1904 A; 398 C; 552 G; 725 T; 0 U; 0 Other;
XX
XX
XX Query Match 11.9%; Score 176.6; DB 3; Length 3579;
XX Best Local Similarity 46.3%; Pred. No. 1.6e-16;
XX Matches 661; Conservative 0; Mismatches 759; Indels 9; Gaps 2;
XX
XX
XX 36 TGCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 95
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 744 TGATTAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 803
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX
XX 96 AAGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 804 TAAGAAACATTTCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 863
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX
XX 156 AGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 215
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 864 AATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX
XX 216 AAGCATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 275
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 924 TGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 276 AGAACAGTGAAGTCTGCTAAGAAAAAGTTACAAGACGCAAGCAATTTAGACAAAG 335
DB 984 AAAAACTGATTAATATATAGCAGATATTAATAAAGAAAGAAAGAAACAAA 1043
QY 336 GAGACGTCTAAGAAAAAGTTGCAAGAACACAAAGCGATTTTGAACAAGAGAGCGTGC 395
DB 1044 AGATGAAAAAGAAAAAATTCATTCACAGCTGTAAAAAGTGTACATTAATTAAGTTGG 1103
QY 396 TAAAGAAAAAGTTGCAAGAACAAAGCGATTTTGAACAAGAGAGCTTCTAAGAAAA 455
DB 1104 AGAAGAAAAAGAAAGCGAAGAAAAAGAACAAAGAGGACAGATGATGAAGAAAGATC 1163
QY 456 GTTGAAGAACAAAGCGATTTTGAACAAGAGAGCGTCTAAGAAAAAGTTGCAAGAAC 515
DB 1164 AGATGATGAAGAAATACAGATGATGAAGAGATACAGATGATGAAGAAAGATCAAGTGA 1223
QY 516 ACAACAAA-----CGCATTTAGAACAAAGAGAGCGTCTAAGAAAAAGTTGCAAGAAC 569
DB 1224 TGAAGAAATACAGGTGATCAAGAAAAACAAAGAAAGAAAGAAAGTGAACGAAAAAAC 1283
QY 570 ACAAAAGCGATTTAGAACAAAGAGAGCGTCTAAGAAAAAGTTGCAAGAGCAGCAAGAGA 629
DB 1284 AGAAAAAGCGCAAGAAATTTAGAAAGAACAAAGAGATCAGAAAAAGACAAAGAAAGA 1343
QY 630 TTTAGAACAAAGAGCGCTGATACGAAAAAAATTTAGAAAGAAAAAGAAAGCAATGGAGA 689
DB 1344 ATCAGAAAAAGATTAAGAAAGATCAGAAAAAGTAAAGAAATCAGAAAAAGCAAAAGA 1403
QY 690 TATATTAGCAGAGATTTATATATGTCGTTAGAAATCCAGCTATAGAACTTCATCAGAGA 749
DB 1404 AAAAACTGAAAGAAATGAAGAAAAAATCTGAAGCGAAAAAGAAAGAAAGTACAAAAA 1463
QY 750 AAATGAACGTGATTTATATATACCATATCTTCTTAACTCAGAGACAAAGAGGGA 809
DB 1464 AGAAACAGATGTATGAAAAAAGAAAAAGGAGATATGAGAGGGAACGATGATGA 1523
QY 810 TAGTAGAGTTCCAGAGAAATCTATATATAGAAAAAATAATAGAAATCTATTAACAC 869
DB 1524 AGAAGCAAAAGAAAAAGAAAGAAAGCAAGCAAGAAACAAAGTGAAGAAAAAGAAACAGA 1583
QY 870 AAATGTTGAAGACGAAGGATATACATTAAGGACATCTTGAAGAAAGAAAGATGTTTC 929
DB 1584 AAAAAAGCAAGAGGGAACAGATTTATGAAGAAAGTACAGATGATTCAGACAAAGATGAAGA 1643
QY 930 AATTAACCAAGAACAAAGAAAGATTAATCTGTCATATCAAAATCATACATTAGAGAC 989
DB 1644 AACAAAGATGAAAGAAAAAGAAACAGAAAGAGACGAAGAAAGAAAGCAAGACGAAAA 1703
QY 990 AGTAAATATTTGATGTTAATGATTTTCAAAATTAAGTATGAGATGAAGAAATAGTGC 1049
DB 1704 AAAAAACAGAGTGAAGAAAAAGAAAAAGAAAAAGCAAGAGGGAACAGATTAAGAGA 1763
QY 1050 TGAATATGACGATTCATTAATAGATGAAGAGAGATGATGAAGACTTGAAGCAATTTAA 1109
DB 1764 AGATACAGATGATTCAGACAAAGATGTAGAAACAGAAAGTGAAGAAACAGACGAGAAGA 1823
QY 1110 GCCTATTTGCAATATGACAAATTTCCAAAGTGAAGAAAAATATGCAATTTATTAAGACT 1169
DB 1824 CAAAGAAAGAAAAAGAAAGGGAACAGATGATGAAGAAAGCAAAAGTGAAGAAACAGACCT 1883
QY 1170 AGAAGTTTGAATGAGAAAAATGAATTTTAAAGTGTATTTTGAAGAAAGTGAAGAAATTC 1229
DB 1884 AGACGCAAGAAAGAGACGGAAGAGATTAAGAAAGCAAGCAAAAGAAAGAAAGCAAGAA 1943
QY 1230 ATCAGAGAGATTTATCGAAGAAAAATTAAGAAAAAGAAATATGAAAAAGCAAGAGA 1289
DB 1944 AGATGCAAAAGAAAGATGACAAAGAAAAAGCAAGAAAGAAAGTGAACAAAGAAATATCAAGA 2003
QY 1290 TAAATATTTTAAACCAATGATTAAGATTGTTGATGATGATATTTAAAAATATATAAAA 1349
DB 2004 AGATGAC---AAAGAAATGACAAAGAAAGATGACAAAGAAAGAAAGCAAAAGAAATACAA 2060

```

```

QY 1350 TGATTAAGCAGGTTAATAGAAAAAGAAAAATTCATTAATCATTTGATATATTTGA 1409
DB 2061 AGAAAAAGACAAAGAGATTAACAAAGAAAAAGCAAAAGAGATGACAAAGAAAAAGACAA 2120
QY 1410 CGGAGACATGAAATTTTACGATCGTGATGAGTATCTGTAAGTATTA 1458
DB 2121 AGAAGATGACAAAGAAAAAGCAAAAGAAAGATTAACAAAGAAAAAGCAAA 2169

RESULT 15
AAA70212
ID AAA70212 standard; DNA; 1998 BP.
XX
AC AAA70212;
XX
DE 07-NOV-2000 (first entry)
XX
XX Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:345.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX antimalarial; malaria; protozoicide; infection; insecticide; ds.
XX
OS Plasmodium falciparum.
XX
PN WO20025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US026796.
XX
PR 05-NOV-1998; 98US-0107131P.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI, 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
XX Plasmodium falciparum, useful as antimalarial vaccines and in the
XX diagnosis of P.falciparum infection.
XX
PS Disclosure: Page 537; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II). (I) and
CC (II) are useful for the development of vaccines against P. falciparum
CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
CC immunogens comprising the sequences of (I), are useful in the detection
CC of infection with P. falciparum. Furthermore, (I) (especially when they
CC are rifins or secreted or membrane proteins) can aid the identification
CC of drugs to treat or prevent P. falciparum infection, or they can be used
CC to identify drug resistance in P. falciparum. Sequencing of the
CC Plasmodium chromosome 2 and the subsequent identification of proteins
CC encoded by it will help to expand our understanding of parasite biology,
CC a process hampered by the complexity of the parasitic lifecycle, and
CC provide new targets for vaccine and drug development. Parasite resistance
CC to drugs and mosquito resistance to insecticides have led to a resurgence
CC of malaria in many parts of the world, and there is a pressing need for
CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
CC represent nucleotide and protein sequences given in the present
CC invention, but which are not specifically mentioned within the
CC specification
XX
SQ Sequence 1998 BP; 1127 A; 141 C; 338 G; 392 T; 0 U; 0 Other;
XX
Query Match 11.5%; Score 170.6; DB 3; Length 1998;
Best Local Similarity 45.7%; Pred. No. 1.2e-15;

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 23:09:13 ; Search time 7825 Seconds

(Without alignments)
10765.750 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1482

Sequence: 1 CAAGACACCAAGCCGATCT.....AATATTTATGAACTATAA 1482

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_in:*

3: gb_env:*

4: gb_on:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pi:*

9: gb_to:*

10: gb_gw:*

11: gb_gy:*

12: gb_un:*

13: gb_vt:*

14: gb_ncg:*

15: gb_d1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1482	100.0	1482	6	ARI62970 Sequence
2	1482	100.0	1482	6	ARI78259 Sequence
3	1482	100.0	1496	2	PFLSA13
4	1480.4	99.9	1482	6	ARI62974 Sequence
5	1480.4	99.9	1482	6	ARI78263 Sequence
6	1446	97.6	1493	6	A28743
7	1446	97.6	1493	6	ARI62967 Sequence
8	1446	97.6	1493	6	ARI78256 Sequence
9	1442	97.3	5970	2	PFLSA1G
10	1406.8	94.9	253001	2	AE014834
11	950	64.1	950	6	A28740
12	950	64.1	950	6	ARI62961 Sequence
13	950	64.1	950	6	ARI78250 Sequence
14	843	56.9	843	2	PFLSA1C
15	841.4	56.8	843	2	PFLSA1B
16	841.4	56.8	843	2	PFLSA1D
17	841.4	56.8	843	2	PFLSA1H
18	841.4	56.8	843	2	PFLSA1I

19	839.8	56.7	843	2	PFLSA1E	L40887 Plasmodium
20	839.8	56.7	843	2	PFLSA1F	L40888 Plasmodium
21	839.8	56.7	843	2	PFLSA1G	L40893 Plasmodium
22	839.8	56.7	843	2	PFLSA1H	L40834 Plasmodium
23	838.2	56.6	843	2	PFLSA1I	L40835 Plasmodium
24	836.6	56.5	843	2	PFLSA1J	L40947 Plasmodium
25	836.6	56.5	843	2	PFLSA1K	L40886 Plasmodium
26	836.6	56.5	843	2	PFLSA1L	L40886 Plasmodium
27	836.6	56.5	843	2	PFLSA1M	L40837 Plasmodium
28	835	56.3	843	2	PFLSA1N	L40910 Plasmodium
29	835	56.3	843	2	PFLSA1O	L40892 Plasmodium
30	833.4	56.2	843	2	PFLSA1P	L40908 Plasmodium
31	833.4	56.2	843	2	PFLSA1Q	L40909 Plasmodium
32	791.8	53.4	795	2	AY751501	AF246996 Plasmodi
33	503.4	34.0	1368	11	AY751501	AY751501 Synthetic
34	472.8	31.9	988	6	PFLSA1S	Z30319 P.falciparu
35	472.8	31.9	988	6	A28742	A28742 cDNA for LS
36	472.8	31.9	988	6	ARI62963	ARI62963 Sequence
37	472.8	31.9	988	6	ARI78252	ARI78252 Sequence
38	470.8	31.8	954	6	ARI62966	ARI62966 Sequence
39	470.8	31.8	954	6	ARI78255	ARI78255 Sequence
40	280.6	18.9	198224	9	AC149597	AC149597 Mus muscu
41	274.4	18.5	208332	9	AC122227	AC122227 Mus muscu
42	252.4	17.0	172716	14	AC115383	AC115383 Rattus no
43	250.8	16.9	184015	9	AC144849	AC144849 Mus muscu
44	237.4	16.0	161482	14	AC165336	AC165336 Mus muscu
45	225.8	15.2	2069	6	E10125	E10125 DNA encodin

ALIGNMENTS

RESULT 1

LOCUS ARI62970 1482 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 41 from patent US 6270771.

ACCESSION ARI62970

VERSION ARI62970.1 GI:16233435

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1482)

AUTHORS Guerin-Marchand,C. and Druilhe,P.

TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes

JOURNAL Patent: US 6270771-A 41 07-AUG-2001.

FEATURES

source Location/Qualifiers

1..1482

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 1482; DB 6; Length 1482;

Best Local Similarity 100.0%; Pred. No. 1,3e-158;

Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGACACCAAGCCGATCTAGAACAAAGAGAGCGTCTTAAGAAAGTTGCAAGAACAA 60

DB 1 CAAGACACCAAGCCGATCTAGAACAAAGAGAGCGTCTTAAGAAAGTTGCAAGAACAA 60

QY 61 CAAGCGATTTAGAACAAAGATGATCTTCTTAAGAAAGTTGCAAGAGCGCAAGCGAT 120

DB 61 CAAGCGATTTAGAACAAAGATGATCTTCTTAAGAAAGTTGCAAGAGCGCAAGCGAT 120

QY 121 TTAGACACAGAGACTTGTCTTAAGAAAGTTGCAAGAACAAAGCGATCTTGAACAA 180

DB 121 TTAGACACAGAGACTTGTCTTAAAGAAAGTTGCAAGAACAAAGCGATCTTGAACAA 180

QY 181 GAGAGACGTGCTTAAGAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCT 240

DB 181 GAGAGACGTGCTTAAGAAAGTTGCAAGAACAAAGCGATTTTGAACAGAGAGAGCT 240

241 GCTAAGAAAAGTTCAGAGACAAAGCGATTTAGAACAGATAGACTTCTTAAGNA 300
241 GCTAAGAAAAGTTCAGAGACAAAGCGATTTAGAACAGATAGACTTCTTAAGNA 300
301 AAGTTACAGAGACGAGCAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCA 360
301 AAGTTACAGAGACGAGCAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCA 360
361 GAAACAACAAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAA 420
361 GAAACAACAAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAA 420
421 AGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
421 AGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
481 GAAACAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
481 GAAACAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAG 540
541 AGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTA 600
541 AGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTA 600
601 AAGAAAAGTTGCAAGAGAGCGATTTAGAACAAAGAGAGCGTCTTAAGAAAAGTT 660
601 AAGAAAAGTTGCAAGAGAGCGATTTAGAACAAAGAGAGCGTCTTAAGAAAAGTT 660
661 AATTTAGAAAAGAAAAGAGACATGAGATATATTAGACAGAGATTTATATGCTGTTA 720
661 AATTTAGAAAAGAAAAGAGACATGAGATATATTAGACAGAGATTTATATGCTGTTA 720
721 GAAATACAGCATATGAACTTCATCAGAAAATGAAACGAGATTTATATACACATCA 780
721 GAAATACAGCATATGAACTTCATCAGAAAATGAAACGAGATTTATATACACATCA 780
781 TCTTTCTTACCTCAGAGCAACAGAGGATATAGAGATTTCCAGAAAATATCTATA 840
781 TCTTTCTTACCTCAGAGCAACAGAGGATATAGAGATTTCCAGAAAATATCTATA 840
841 GAAAAAACAATAGAGATCTATTACACAAATGTTGAAGACGAAAGGATTTACATTA 900
841 GAAAAAACAATAGAGATCTATTACACAAATGTTGAAGACGAAAGGATTTACATTA 900
901 GGAACATCTTGAAGAAAAGAGTGTGTAATTAACAGAACAAAGAAATTAATCT 960
901 GGAACATCTTGAAGAAAAGAGTGTGTAATTAACAGAACAAAGAAATTAATCT 960
961 GCTGACATACAAAATCATATTTAGAGACATTAATATTTCTGATGTTAATGATTTCA 1020
961 GCTGACATACAAAATCATATTTAGAGACATTAATATTTCTGATGTTAATGATTTCA 1020
1021 AATAGTAAATAGAGATTAATAGTGTGTAATTAATGATGATTTCAATTAATGATGA 1080
1021 AATAGTAAATAGAGATTAATAGTGTGTAATTAATGATGATTTCAATTAATGATGA 1080
1081 GAAAGTGAAGAGACCTTAACAAATTAAGCCATATGTCCAATATGACAAATTTCCA 1140
1081 GAAAGTGAAGAGACCTTAACAAATTAAGCCATATGTCCAATATGACAAATTTCCA 1140
1141 GAAAGAAAACATAGAAATTTATTAAGAACATAGAAATTTATGAGAAAATGATTA 1200
1141 GAAAGAAAACATAGAAATTTATTAAGAACATAGAAATTTATGAGAAAATGATTA 1200
1201 GATGATTTAGAGAGAAATTAAGAAAATCATGAGAAATTTATCTGAGAAAATTA 1260
1201 GATGATTTAGAGAGAAATTAAGAAAATCATGAGAAATTTATCTGAGAAAATTA 1260
1261 AAGAAAAGAAATTAAGAAAACAAAGGATTAATTTTAACCAATATGATTAAGTTG 1320
1261 AAGAAAAGAAATTAAGAAAACAAAGGATTAATTTTAACCAATATGATTAAGTTG 1320
1321 TATGATGACATATTTAAAAATATATAATGATTAAGCAGTTAATTAAGAAAAGGAAA 1380

1321 TATGATGACATATTTAAAAATATATAATGATTAAGCAGTTAATTAAGAAAAGGAAA 1380
1381 TTCAATTAATGATTTTCAATATATTTGACGAGACATGAAATTTTACAGATCGTGAT 1440
1381 TTCAATTAATGATTTTCAATATATTTGACGAGACATGAAATTTTACAGATCGTGAT 1440
1441 GAGTTATCTGAAGATATTAATTAATTTTATGAAACTATTA 1482
1441 GAGTTATCTGAAGATATTAATTAATTTTATGAAACTATTA 1482

RESULT 2
AR178259 1482 bp DNA linear PAT 20-APR-2002
LOCUS AR178259
DEFINITION Sequence 41 from patent US 6319502.
ACCESSION AR178259
VERSION AR178259.1 GI:20219397
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1482)
AUTHORS Guerin-Marchand, C. and Druille, P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6319502-A 41 20-NOV-2001;
FEATURES
location/Qualifiers
1..1482
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 1482; DB 6; Length 1482;
Best Local Similarity 100.0%; Pred. No. 1.3e-158;
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAAGAACAAACAAAGCATCTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGACAA 60
1 CAAGAACAAACAAAGCATCTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGACAA 60
61 CAAGCGATTTAGAACAGATAGACTTCTTAAGAAAAGTTTCAAGAGACGAAAGCGAT 120
61 CAAGCGATTTAGAACAGATAGACTTCTTAAGAAAAGTTTCAAGAGACGAAAGCGAT 120
121 TTAGAACAGAGAGAGCTTGTCTTAAGAAAAGTTGCAAGAACAAAGGATCTAGAACAA 180
121 TTAGAACAGAGAGAGCTTGTCTTAAGAAAAGTTGCAAGAACAAAGGATCTAGAACAA 180
181 GAGAGACGCTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGT 240
181 GAGAGACGCTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGT 240
241 GCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATAGACTTCTTAAGAAA 300
241 GCTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATAGACTTCTTAAGAAA 300
301 AAGTTACAGAGACGAGCAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCA 360
301 AAGTTACAGAGACGAGCAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCA 360
361 GAAACAACAAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAA 420
361 GAAACAACAAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAA 420
421 AGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
421 AGCGATTTAGAACAGAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
481 GAAACAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGGATTTAGAACAGAG 540
481 GAAACAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGGATTTAGAACAGAG 540
481 GAAACAAGAGAGCGTCTTAAGAAAAGTTGCAAGAACAAAGGATTTAGAACAGAG 540

181 GAGAGACGTGCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGACGT 240
181 GAGAGACGTGCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGACGT 240
181 GAGAGACGTGCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGACGT 240
241 GCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGACGT 300
241 GCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGACGT 300
301 AAGTTAACAAGACCAAAAGCGATTTAGAACAGAGACGT 360
301 AAGTTAACAAGACCAAAAGCGATTTAGAACAGAGACGT 360
301 AAGTTAACAAGACCAAAAGCGATTTAGAACAGAGACGT 360
361 GAACAAACAAAGCGATTTAGAACAGAGACGT 420
361 GAACAAACAAAGCGATTTAGAACAGAGACGT 420
421 AGCGATTTAGAACAGAGACGT 480
421 AGCGATTTAGAACAGAGACGT 480
421 AGCGATTTAGAACAGAGACGT 480
481 GAACAAACAAAGCGATTTAGAACAGAGACGT 540
481 GAACAAACAAAGCGATTTAGAACAGAGACGT 540
481 GAACAAACAAAGCGATTTAGAACAGAGACGT 540
541 AGCGATTTAGAACAGAGACGT 600
541 AGCGATTTAGAACAGAGACGT 600
541 AGCGATTTAGAACAGAGACGT 600
601 AAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGACGT 660
601 AAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGACGT 660
601 AAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGACGT 660
661 AATTTAGAAAAGAAAAGGAGATGAGATATTTAGAACAGAGACGT 720
661 AATTTAGAAAAGAAAAGGAGATGAGATATTTAGAACAGAGACGT 720
661 AATTTAGAAAAGAAAAGGAGATGAGATATTTAGAACAGAGACGT 720
721 GAATATCACAGTATAGAACTTCATCAGAAAATGAAAGTATTTATATACATCA 780
721 GAATATCACAGTATAGAACTTCATCAGAAAATGAAAGTATTTATATACATCA 780
721 GAATATCACAGTATAGAACTTCATCAGAAAATGAAAGTATTTATATACATCA 780
781 TCTTCTTTACCTCAGAGCAACAGAGGAGATAGAGATTTCCAGGAAATATCTATATA 840
781 TCTTCTTTACCTCAGAGCAACAGAGGAGATAGAGATTTCCAGGAAATATCTATATA 840
781 TCTTCTTTACCTCAGAGCAACAGAGGAGATAGAGATTTCCAGGAAATATCTATATA 840
841 GAAAGAAAAGGAGATGAGATATTTAGAACAGAGACGT 900
841 GAAAGAAAAGGAGATGAGATATTTAGAACAGAGACGT 900
841 GAAAGAAAAGGAGATGAGATATTTAGAACAGAGACGT 900
901 GGAACATCTTGAAGAAAAGAGTGTTCATTTAAACCAAGAACAAAGATTAATCT 960
901 GGAACATCTTGAAGAAAAGAGTGTTCATTTAAACCAAGAACAAAGATTAATCT 960
901 GGAACATCTTGAAGAAAAGAGTGTTCATTTAAACCAAGAACAAAGATTAATCT 960
961 GCTGACATCAAAATATCATTTAGAGACAGTAAATTTCTGATGTTATATTTTCA 1020
961 GCTGACATCAAAATATCATTTAGAGACAGTAAATTTCTGATGTTATATTTTCA 1020
961 GCTGACATCAAAATATCATTTAGAGACAGTAAATTTCTGATGTTATATTTTCA 1020
1021 AATTAAGTAATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1080
1021 AATTAAGTAATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1080
1021 AATTAAGTAATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1080
1081 GAAGATGATGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1140
1081 GAAGATGATGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1140
1081 GAAGATGATGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1140
1141 GAAGATGATGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1200
1141 GAAGATGATGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1200
1141 GAAGATGATGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1200
1201 GATGATTTAGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1260
1201 GATGATTTAGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1260
1201 GATGATTTAGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1260
1261 AAGGATTTAGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1320
1261 AAGGATTTAGAGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1320

1261 AAAAGAAAAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1320
1261 AAAAGAAAAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1320
1261 AAAAGAAAAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1320
1321 TATGATGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1380
1321 TATGATGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1380
1321 TATGATGAGATTTAGAGACATTTAGAGATTTAGAGATTTAGAGATTTAGAGATTTAGAGAT 1380
1381 TTTATTAAGTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1440
1381 TTTATTAAGTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1440
1381 TTTATTAAGTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1440
1441 GAGTTATCTGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1482
1441 GAGTTATCTGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1482
1441 GAGTTATCTGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1482

RESULT 4
AR162974
LOCUS AR162974 1482 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 45 from patent US 6270771.
ACCESSION AR162974
VERSION AR162974.1 GI:16233440
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1482)
Guertin-Marchand, C. and Druille, P.
Peptide sequences specific for the hepatic stages of P. falciparum
bearing epitopes capable of stimulating the T lymphocytes
Patent: US 6270771-A 45 07-AUG-2001
LOCATION/Qualifiers
1..1482
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 99.9%; Score 1480.4; DB 6; Length 1482;
Best Local Similarity 99.9%; Pred. No. 2e-158;
Matches 1481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CAAGAACAAACAAAGCGATCTTGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 60
1 CAAGAACAAACAAAGCGATCTTGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 60
1 CAAGAACAAACAAAGCGATCTTGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 60
61 CAAGCGATTTAGAACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 120
61 CAAGCGATTTAGAACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 120
61 CAAGCGATTTAGAACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 120
121 TTAGAACAGAGAGACGTCTTGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 180
121 TTAGAACAGAGAGACGTCTTGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 180
121 TTAGAACAGAGAGACGTCTTGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 180
121 TTAGAACAGAGAGACGTCTTGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 180
181 GAGAGACGTGCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGAGACGT 240
181 GAGAGACGTGCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGAGACGT 240
181 GAGAGACGTGCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGAGACGT 240
241 GCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGATGAGATGAGATGAGATGAGAT 300
241 GCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGATGAGATGAGATGAGATGAGAT 300
241 GCTAAAGAAAAGTTGCAGAACAAACGATTTAGAACAGAGATGAGATGAGATGAGATGAGAT 300
301 AAGTTAACAAGACCAAAAGCGATTTAGAACAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 360
301 AAGTTAACAAGACCAAAAGCGATTTAGAACAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 360
301 AAGTTAACAAGACCAAAAGCGATTTAGAACAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 360
301 AAGTTAACAAGACCAAAAGCGATTTAGAACAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 360
361 GAACAAACAAAGCGATTTAGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 420
361 GAACAAACAAAGCGATTTAGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 420
361 GAACAAACAAAGCGATTTAGAACAGAGAGACGTCTTAAAGAAAAGTTGCAGAACAA 420
421 AGCGATTTAGAACAGAGACGTCTTAAAGAAAAGTTGCAGAACAAACGATTTA 480
421 AGCGATTTAGAACAGAGACGTCTTAAAGAAAAGTTGCAGAACAAACGATTTA 480
421 AGCGATTTAGAACAGAGACGTCTTAAAGAAAAGTTGCAGAACAAACGATTTA 480

QY 781 TCTTCTTACCTCAGACAAACAGAGGAAATAGTAGAGATTCACAGAAATATCTATATA 840
DB 781 TCTTCTTACCTCAGACAAACAGAGGAAATAGTAGAGATTCACAGAAATATCTATATA 840
QY 841 GAAAAAACAATAGAGAACTATTAACAACAAATGTTGAAGAGACGAAGGATATACATATA 900
DB 841 GAAAAAACAATAGAGAACTATTAACAACAAATGTTGAAGAGACGAAGGATATACATATA 900
QY 901 GGACATCTTGAAGAAAAGATGCTTCAATTAACCCAGAACCAAAAAGAGATTAATCT 960
DB 901 GGACATCTTGAAGAAAAGATGCTTCAATTAACCCAGAACCAAAAAGAGATTAATCT 960
QY 961 GCTGACATACAAATCATCATTAAGACAGATAAATTTCTGATGTTAATGATTTTCAA 1020
DB 961 GCTGACATACAAATCATCATTAAGACAGATAAATTTCTGATGTTAATGATTTTCAA 1020
QY 1021 ATTAAGTATAGAGATGAAAATTAAGTCTGAATTAAGACATTCATTAATAGTAGAAGA 1080
DB 1021 ATTAAGTATAGAGATGAAAATTAAGTCTGAATTAAGACATTCATTAATAGTAGAAGA 1080
QY 1081 GAAGATGATGAAGACTTAGACGAATTTAAGCCTATTGTCGAATATGACAAATTTCCAGAT 1140
DB 1081 GAAGATGATGAAGACTTAGACGAATTTAAGCCTATTGTCGAATATGACAAATTTCCAGAT 1140
QY 1141 GAAGAAAACATAGAAATTTAATAAGAACTAGAAAGATTTGATAGAGAAAATGAATAATTA 1200
DB 1141 GAAGAAAACATAGAAATTTAATAAGAACTAGAAAGATTTGATAGAGAAAATGAATAATTA 1200
QY 1201 GATGATTTAGATGAAGAAATAGAAAATCATCAGAAAGATTTATCTGAAGAAAATTAATA 1260
DB 1201 GATGATTTAGATGAAGAAATAGAAAATCATCAGAAAGATTTATCTGAAGAAAATTAATA 1260
QY 1261 AAAGGAAAAGAAATAGAAAACCAAGGATTAATTTTAAACCAATGATAAAAGTTTG 1320
DB 1261 AAAGGAAAAGAAATAGAAAACCAAGGATTAATTTTAAACCAATGATAAAAGTTTG 1320
QY 1321 TATGATGACATATTTAAAAATATATAAATGATTAAGCAGGTTATTAAGAAAAGAAAA 1380
DB 1321 TATGATGACATATTTAAAAATATATAAATGATTAAGCAGGTTATTAAGAAAAGAAAA 1380
QY 1381 TTCAATAAATCATTTGTTCAATATTTTGAACGAGACATGAATTTTACAGATCGTGAT 1440
DB 1381 TTCAATAAATCATTTGTTCAATATTTTGAACGAGACATGAATTTTACAGATCGTGAT 1440
QY 1441 GAGTTATCTGAAGATTAATTAATTTTATGAACCTATA 1482
DB 1441 GAGTTATCTGAAGATTAATTAATTTTATGAACCTATA 1482

RESULT 6
A28743 1493 bp DNA linear PAT 04-JUN-1995
LOCUS A28743
DEFINITION CDNA for USA gene 3' (partial).
ACCESSION A28743
VERSION A28743.1 GI:1247516
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1493)
AUTHORS
TITLE LIVER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING
JOURNAL EPIPTOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES
FEATURES Patent: WO 9213884-A 32 20-AUG-1992;
SOURCE Location/Qualifiers
1..1493
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 97.6%; Score 1446; DB 6; Length 1493;
Best Local Similarity 99.8%; Pred. No. 1.5e-154;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
QY 1 CAAGAACAAACAAGCGATCTTGAAACAAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAA 60
DB 1 CAAGAACAAACAAGCGATCTTGAAACAAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAA 60
QY 61 CAAGCGATTTTGAACAAGATAGACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTTGAACAA 120
DB 61 CAAGCGATTTTGAACAAGATAGACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTTGAACAA 120
QY 121 TTAGAACAGAGAGACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTTGAACAA 180
DB 121 TTAGAACAGAGAGACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTTGAACAA 179
QY 181 GAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGAGAGCT 240
DB 181 GAGAGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGAGAGCT 239
QY 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGAGAGCTTGAACAA 300
DB 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGAGAGCTTGAACAA 299
QY 301 AAGTTACAGAGAGACCAAGCGATTTTGAACAGAGAGAGCTGCTTAAAGAAAAGTTGCA 360
DB 301 AAGTTACAGAGAGACCAAGCGATTTTGAACAGAGAGAGCTGCTTAAAGAAAAGTTGCA 359
QY 361 GAACAAACAAGCGATTTTGAACAAAGAGAGAGCTGCTTAAAGAAAAGTTGCAAGAACAA 420
DB 361 GAACAAACAAGCGATTTTGAACAAAGAGAGAGAGCTGCTTAAAGAAAAGTTGCAAGAACAA 418
QY 421 AGCGATTTTGAACAAGAGAGAGCTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGA 480
DB 421 AGCGATTTTGAACAAGAGAGAGCTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGA 478
QY 479 GAACAG 538
DB 479 GAACAG 538
QY 541 AGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGAGAGAGCTGCT 600
DB 541 AGACGTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTTGAACAAAGAGAGAGAGCTGCT 597
QY 597 AAAGAAAAGTTGCAAGAGACCAAGAGATTTTGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 597 AAAGAAAAGTTGCAAGAGACCAAGAGATTTTGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAG 657
QY 661 AATTTAGAAAGAAAAGAAAAGAAATGAGATATTTAGCAAGAGATTTATATGTCGTTTA 720
DB 661 AATTTAGAAAGAAAAGAAAAGAAATGAGATATTTAGCAAGAGATTTATATGTCGTTTA 717
QY 721 GAATATCAAGCTATAGAACTTCATCAGAAAATGAACGTGATATTTATACCAATCA 780
DB 721 GAATATCAAGCTATAGAACTTCATCAGAAAATGAACGTGATATTTATACCAATCA 777
QY 778 TCTTCTTACCTCAGAGCAACAGAGGAAATGATAGAGATTCAGAGAAATATCTATATA 837
DB 778 TCTTCTTACCTCAGAGCAACAGAGGAAATGATAGAGATTCAGAGAAATATCTATATA 837
QY 841 GAAAAAACAATAGAGAACTATTAACAACAAATGTTGAAGAGACGAAGGATATACATATA 900
DB 841 GAAAAAACAATAGAGAACTATTAACAACAAATGTTGAAGAGACGAAGGATATACATATA 897
QY 901 GGACATCTTGAAGAAAAGATGCTTCAATTAACCCAGAACCAAAAAGAGATTAATCT 960
DB 901 GGACATCTTGAAGAAAAGATGCTTCAATTAACCCAGAACCAAAAAGAGATTAATCT 957
QY 961 GCTGACATACAAATCATCATTAAGACAGATAAATTTCTGATGTTAATGATTTTCAA 1020
DB 961 GCTGACATACAAATCATCATTAAGACAGATAAATTTCTGATGTTAATGATTTTCAA 1017
QY 1021 ATTAAGTATAGAGATGAAAATTAAGTCTGAATTAAGACATTCATTAATAGTAGAAGA 1080

Db	1018	ATTAAGTAAGTATGAGATGAAATAAGCTGAATATGCGATTCAATTATATGATGAAGAA	1077
Oy	1081	GAAGTATGTAAGAAGCTTAGACGAATTTAAAGCTATTGCGACAATATGCAATTTCCAAAGT	1140
Db	1078	GAAGTATGTAAGACTTAGACGAATTTAAAGCTATTGCGACAATATGCAATTTCCAAAGT	1137
Oy	1141	GAAGAAACATAGGAATTTATTAAGAACTAGAGAATTTGATAGAGAAAAATGAAAAATTTA	1200
Db	1138	GAAGAAACATAGGAATTTATTAAGAACTAGAGAATTTGATAGAGAAAAATGAAAAATTTA	1197
Oy	1201	GATGATTTAGATGAAGGAATGAAAAATCATCAGAGAAATTTATCTGAAGAAAAATTTAAA	1260
Db	1198	GATGATTTAGATGAAGGAATGAAAAATCATCAGAGAAATTTATCTGAAGAAAAATTTAAA	1257
Oy	1261	AAAGGAAGAATATGAAAAAACAAGAGTATATAATTTTAAACCAATGATTAAGATTG	1320
Db	1258	AAAGGAAGAATATGAAAAAACAAGAGTATATAATTTTAAACCAATGATTAAGATTG	1317
Oy	1321	TATGATGAGCATATTTAAAAAATATTAATAATGATAGCAGTTAATTAAGAAAAAGAAAA	1380
Db	1318	TATGATGAGCATATTTAAAAAATATTAATAATGATAGCAGTTAATTAAGAAAAAGAAAA	1377
Oy	1391	TTCATTAATCATTTGTTTCATATATTTGACGAGACAAATGAATTTTACAGTGTGAT	1440
Db	1378	TTCATTAATCATTTGTTTCATATATTTGACGAGACAAATGAATTTTACAGTGTGAT	1437
Oy	1441	GAGTTATCTGAAGATATACTTAATTTTATGAACATATA	1482
Db	1438	GAGTTATCTGAAGATATACTTAATTTTATGAACATATA	1479

[illegible]

OY	241	GCTAAAGAAAAGTTGGCAAGAACAA	CAAAAGGATTTAGAACAA	GATGACGTTGGCTAAAGAA	300
Db	240	GCTAAAGAAAAGTTGGCAAGAACAA	CAAAAGGATTTAGAACAA	GATGACGTTGGCTAAAGAA	299
OY	301	AAGTTACAAGAGCAGCAAAAGCGATTTAGAA	CAAGAGAGAGCGTGTCTAAAGAAAGTTGCCAA		360
Db	300	AAGTTACAAGAGCAGCAAAAGCGATTTAGAA	CAAGAGAGAGCGTGTCTAAAGAAAGTTGCCAA		359
OY	361	GAACAACAAAAGCGATTTAGAACAA	GAGAGACGTGCTAAAGAAAAGTTGCCAAAGAACAA	CAAA	420
Db	360	GAACAACAAAAGCGATTTAGAACAA	GAGAGACGTGCT - AAGAAAAGTTGCCAAAGAACAA	CAAA	418
OY	421	AGCGATTTAGAACAAAGAGACCTTGGCTAAAGAAAAGTTGCAAGAACAA	CAAAAGCGATTTA		480
Db	419	AGCGATTTAGAACAAAGAGAGACCTTGGCTAAAGAAAAGTTGCAAGAACAA	CAAAAGCGATTTA		478
OY	481	GAACAAAGAGACGTGCTAAAGAAAAGTTGCCAAAGAACAA	CAAAAGCGATTTAGAACAA	GAG	540
Db	479	GAACAAAGAGACGTGCTAAAGAAAAGTTGCCAAAGAACAA	CAAAAGCGATTTAGAACAA	GAG	538
OY	541	AGACGTGCTAAAGAAAAGTTGCCAAAGAACAA	CAAAAGCGATTTAGAACAA	GAGAGACGTGCT	600
Db	539	AGACGTGCT - AAGAAAAGTTGCCAAAGAACAA	CAAAAGCGATTTAGAACAA	GAGAGACGTGCT	597
OY	601	AAAGAAAAGTTGCCAAAGACGACCAAAAGATTTAGAACAA	AGAAAGCGTGTATACGAAAAAA		660
Db	598	AAAGAAAAGTTGCCAAAGACGACCAAAAGATTTAGAACAA	AGAAAGCGTGTATACGAAAAAA		657
OY	661	AAATTAGAAAGAAAAAGGAACATGAGATATATTAGACAGAGATTTATATGCGCTTTA			720
Db	658	AAATTAGAAAGAAAAAGGAACATGAGATATATTAGACAGAGATTTATATGCGCTTTA			717
OY	721	GAATATACCAAGCTATAGAACTTCCATCGAAAAATGAACGTGATATTATATCAACATCAA			780
Db	718	GAATATACCAAGCTATAGAACTTCCATCGAAAAATGAACGTGATATTATATCAACATCAA			777
OY	781	TCTTCTTTACTCTGAGACACAGAGGGAATAGTAGAATTTCCAGGAATTTCTATATATA			840
Db	778	TCTTCTTTACTCTGAGACACAGAGGGAATAGTAGAATTTCCAGGAATTTCTATATATA			837
OY	841	GAAAAAACAATATAGAGATCTATATACAA	CAAAAGTTGAAAGGACGAAGGATATCATATAA		900
Db	838	GAAAAAACAATATAGAGATCTATATACAA	CAAAAGTTGAAAGGACGAAGGATATCATATAA		897
OY	901	GGACATCTTTGAAAGAAAAGAAAGATGTTCCATTA	AAACACGAAACAAAAAGAAATAAATCT		960
Db	898	GGACATCTTTGAAAGAAAAGAAAGATGTTCCATTA	AAACACGAAACAAAAAGAAATAAATCT		957
OY	961	GCTGACATACAAAATCATACATATAGACAGATAAATTTCTGATGTTAATGATTTTCAA			1020
Db	958	GCTGACATACAAAATCATACATATAGACAGATAAATTTCTGATGTTAATGATTTTCAA			1017
OY	1021	ATTAAGTAAATATAGAGATGAATTAAGTCTGAATATAGACGATTAATTAATGATGAAGAA			1080
Db	1018	ATTAAGTAAATATAGAGATGAATTAAGTCTGAATATAGACGATTAATTAATGATGAAGAA			1077
OY	1081	GAAGATGATGAGACTTAGACGAATTTAGAGCTATTTGCAATATATGACAAATTTCCAAAT			1140
Db	1078	GAAGATGATGAGACTTAGACGAATTTAGAGCTATTTGCAATATATGACAAATTTCCAAAT			1137
OY	1141	GAAGAAAACATAGCAATTTATTAAGAACTAGAAAGTTGATAGAGAAAATGAAAAATTTTA			1200
Db	1138	GAAGAAAACATAGCAATTTATTAAGAACTAGAAAGTTGATAGAGAAAATGAAAAATTTTA			1197
OY	1201	GATGATTTAGATGAAGAAATAGAAAAATCATACAGAAATTAATCTGAAAGAAAAAATATAA			1266
Db	1198	GATGATTTAGATGAAGAAATAGAAAAATCATACAGAAATTAATCTGAAAGAAAAAATATAA			1257
OY	1261	AAAGGAAAGAAATATGAAAAAAACAAAGATATATATTTTAAACCAAAATGATATAAAGTTTG			1320
Db	1258	AAAGGAAAGAAATATGAAAAAAACAAAGATATATATTTTAAACCAAAATGATATAAAGTTTG			1317
OY	1321	TATGTATGACATATTTAAAAAATATATAAAATGATATAGACGTTTAAATTAAGAAAAAGAAAAA			1380

|||||
1318 TATGATGAGCATATTAATAAATATGATTAAGCAGGTTATATTAAGAAAGAAAA 1377
|||
Qy 1381 TTCATTAATATCTTTGTTTATATATTTTACGAGACATGAATTTTACAGATCGTGAT 1440
|||
Db 1378 TTCATTAATATCTTTGTTTATATATTTTACGAGACATGAATTTTACAGATCGTGAT 1437
|||
Qy 1441 GAGTTATCTGAAGATTAATACTAAATTTTATGAACATATA 1482
|||
Db 1438 GAGTTATCTGAAGATTAATACTAAATTTTATGAACATATA 1479
|||
RESULT 8
AR178256 1493 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 38 from patent US 6319502.
DEFINITION AR178256
ACCESSION AR178256
VERSION AR178256.1 GI:20219394
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1493)
AUTHORS Guerlin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
JOURNAL Patent: US 6319502-A 38 20-NOV-2001;
FEATURES
Location/Qualifiers
1..1493
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 97.6%; Score 1446; DB 6; Length 1493;
Best Local Similarity 99.8%; Pred. No. 1.5e-154;
Matches 1479; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
1
1 CAAAGCAACAAGCGATCTAGAACAAAGAGACGCTAAAGAAAAGTTGCAGAACAA 60
Db 1 CAAAGCAACAAGCGATCTAGAACAAAGAGACGCTAAAGAAAAGTTGCAGAACAA 60
|||
Qy 61 CAAAGCGATTTAGAACAAAGATAGACTTGTAAAGAAAAGTTACAGAGACGCAAGCGAT 120
|||
Db 61 CAAAGCGATTTAGAACAAAGATAGACTTGTAAAGAAAAGTTACAGAGACGCAAGCGAT 120
|||
Qy 121 TTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAAGAACACAAAGCGATCTAGAACAA 180
|||
Db 121 TTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAAGAACACAAAGCGATCTAGAACAA 180
|||
Qy 121 TTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAAGAACACAAAGCGATCTAGAACAA 179
|||
Db 121 TTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAAGAACACAAAGCGATCTAGAACAA 179
|||
Qy 181 GAGAGCGTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGACGT 240
|||
Db 181 GAGAGCGTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGACGT 239
|||
Qy 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTCTAAAGAAA 300
|||
Db 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATAGACTTGTCTAAAGAAA 299
|||
Qy 301 AAGTTCAAGAGAGCAAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCA 360
|||
Db 301 AAGTTCAAGAGAGCAAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAAGTTGCA 359
|||
Qy 361 GAAACAACAAGCGATTTAGAACAAAGAGAGACGCTCTAAAGAAAAGTTGCAAGAACAA 420
|||
Db 361 GAAACAACAAGCGATTTAGAACAAAGAGAGACGCTCTAAAGAAAAGTTGCAAGAACAA 418
|||
Qy 421 AGCGATTTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 480
|||
Db 419 AGCGATTTAGAACAAAGAGACTTGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 478
|||
Qy 481 GAAACAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG 540
|||
Db 479 GAAACAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG 538

Qy 541 AGACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCT 600
|||
Db 539 AGACGTCT-AAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCT 597
|||
Qy 601 AAAGAAAAGTTGCAAGAGACGACCAAGAGATTTAGAACAAAGAGAGCTGATTCGAAAAA 660
|||
Db 598 AAAGAAAAGTTGCAAGAGACGACCAAGAGATTTAGAACAAAGAGAGCTGATTCGAAAAA 657
|||
Qy 661 AATTAGAAAAGAAAAGAAATGAGATATATATAGAGAGATTTATATGTCGTTTA 720
|||
Db 658 AATTAGAAAAGAAAAGAAATGAGATATATATAGAGAGATTTATATGTCGTTTA 717
|||
Qy 721 GAAATACAGCTATAGAACTTCCATCAGAAAAATGAAAGCGATATATATACCATCA 780
|||
Db 718 GAAATACAGCTATAGAACTTCCATCAGAAAAATGAAAGCGATATATATACCATCA 777
|||
Qy 781 TCTTTTCTTACCTCAGAGACAAAGAGGAAATGTAGAGATTTCCAGGAAATCTATATA 840
|||
Db 778 TCTTTTCTTACCTCAGAGACAAAGAGGAAATGTAGAGATTTCCAGGAAATCTATATA 837
|||
Qy 841 GAAAAAACAATPAGGAATCTATTACAACAATGTTGAAGGACGAAGGATATACATATA 900
|||
Db 838 GAAAAAACAATPAGGAATCTATTACAACAATGTTGAAGGACGAAGGATATACATATA 897
|||
Qy 901 GGACATCTTGAAGAAAAGATGCTTCAATTAACCAAGAACAAAGAAAGATTAATCT 960
|||
Db 898 GGACATCTTGAAGAAAAGATGCTTCAATTAACCAAGAACAAAGAAAGATTAATCT 957
|||
Qy 961 GCTGACATCAAAATCATACATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCA 1020
|||
Db 958 GCTGACATCAAAATCATACATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTTCA 1017
|||
Qy 1021 ATTAAGTATGAGAGATGAATAAGTCTGAAATGACGATTCATTAATAGATGAAGA 1080
|||
Db 1018 ATTAAGTATGAGAGATGAATAAGTCTGAAATGACGATTCATTAATAGATGAAGA 1077
|||
Qy 1081 GAAAGATGATGAAGACTTATAGCAATTTAAGCTATTTGCAATATGACAAATTTCCAAGAT 1140
|||
Db 1078 GAAAGATGATGAAGACTTATAGCAATTTAAGCTATTTGCAATATGACAAATTTCCAAGAT 1137
|||
Qy 1141 GAAAGAAAATATGAAATTTATTAAGAACTAGAAAGTTGATGAGAAAAATGAATTTA 1200
|||
Db 1138 GAAAGAAAATATGAAATTTATTAAGAACTAGAAAGTTGATGAGAAAAATGAATTTA 1197
|||
Qy 1201 GATGATTTAGATGAAGATATGAATAATCATCAGAAAGATTTCTGAAGAAAATTA 1260
|||
Db 1198 GATGATTTAGATGAAGATATGAATAATCATCAGAAAGATTTCTGAAGAAAATTA 1257
|||
Qy 1261 AAAGAAAAGAAATATGAATAAACAAGATTAATTTTAAACCAATGATTAAGTTTG 1320
|||
Db 1258 AAAGAAAAGAAATATGAATAAACAAGATTAATTTTAAACCAATGATTAAGTTTG 1317
|||
Qy 1321 TATGATGAGCATATTAATAAATATTAAGTAAGCAGTTTAATTAAGAAAAGAAAA 1380
|||
Db 1318 TATGATGAGCATATTAATAAATATTAAGTAAGCAGTTTAATTAAGAAAAGAAAA 1377
|||
Qy 1381 TTCATTAATATCTTTGTTTATATATTTTACGAGACATGAATTTTACAGATCGTGAT 1440
|||
Db 1378 TTCATTAATATCTTTGTTTATATATTTTACGAGACATGAATTTTACAGATCGTGAT 1437
|||
Qy 1441 GAGTTATCTGAAGATTAATACTAAATTTTATGAACATATA 1482
|||
Db 1438 GAGTTATCTGAAGATTAATACTAAATTTTATGAACATATA 1479
|||
RESULT 9
PFLSAIG 5970 bp DNA linear INV 18-APR-2005
LOCUS P.falciparum USA-1 gene for liver stage antigen.
DEFINITION X56203 S75010 S75012 S75014 S75016
ACCESSION X56203.1 GI:9915
VERSION
KEYWORDS liver stage antigen.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)

[illegible]

	Matches	145/1	Conservative	0	Mismatches	25	Indels	0	Gaps	0
QY	1	CAAGAACAA	CAAAAGCGATCTT	GAACCAAGAGAGACGTCCTTAA	GAAGAAAGTTGCAAGAACAA	60				
Db	4327	CAAGAACAA	CAAAAGCGATCTT	GAACCAAGAGAGACGTCCTTAA	GAAGAAAGTTGCAAGAACAA	4386				
QY	61	CAAAAGCATTTT	GAACCAAGAGATCTT	GAACCAAGAGATCTT	GAACCAAGAGATCTT	120				
Db	4387	CAAAAGCATTTT	GAACCAAGAGATCTT	GAACCAAGAGATCTT	GAACCAAGAGATCTT	4446				
QY	121	TTAGAACAA	GAGAGAGACTT	GTCTTAAAGAAAAGTTGCAAGAACAA	CAACCAAGAGATCTT	180				
Db	4447	TTAGAACAA	GAGAGAGACTT	GTCTTAAAGAAAAGTTGCAAGAACAA	CAACCAAGAGATCTT	4506				
QY	181	GAGAGACGTCCTT	AAAGAAAAGTTGCAAGAACAA	CAACCAAGAGATCTT	GAACCAAGAGAGACGT	240				
Db	4507	GAGAGACGTCCTT	AAAGAAAAGTTGCAAGAACAA	CAACCAAGAGATCTT	GAACCAAGAGAGACGT	4566				
QY	241	GCTTAAAGAAAAGTTT	GCAAGAACAAAGGATTTT	GAACCAAGATGACCTT	GTCTTAAAGAA	300				
Db	4567	GCTTAAAGAAAAGTTT	GCAAGAACAAAGGATTTT	GAACCAAGATGACCTT	GTCTTAAAGAA	4626				
QY	301	AAGTTAAAGAGCAGCAAGACGATTTT	GAACCAAGAGAGACGTCCTTAAAGAAAAGTTGCA	360						
Db	4627	AAGTTAAAGAGCAGCAAGACGATTTT	GAACCAAGAGAGACGTCCTTAAAGAAAAGTTGCA	4686						
QY	361	GAACCAACAAACGATTTT	GAACCAAGAGAGACGTCCTTAAAGAAAAGTTGCAAGAACAA	420						
Db	4687	GAACCAACAAACGATTTT	GAACCAAGAGAGACGTCCTTAAAGAAAAGTTGCAAGAACAA	4746						
QY	421	AGCGATTTT	GAACCAAGAGACGTCCTTAAAGAAAAGTTGCAAGAACAAAGGATTTT	480						
Db	4747	AGCGATTTT	GAACCAAGAGACGTCCTTAAAGAAAAGTTGCAAGAACAAAGGATTTT	4806						
QY	481	GAACCAAGAGAGACGTCCTT	AAAGAAAAGTTGCAAGAACAA	CAACCAAGAGATTTT	GAACCAAGAG	540				
Db	4807	GAACCAAGAGAGACGTCCTT	AAAGAAAAGTTGCAAGAACAA	CAACCAAGAGATTTT	GAACCAAGAG	4866				
QY	541	AGACGTCCTT	AAAGAAAAGTTGCAAGAACAA	CAACCAAGAGATTTT	GAACCAAGAGACGTCCT	600				
Db	4867	AGACGTCCTT	AAAGAAAAGTTGCAAGAACAA	CAACCAAGAGATTTT	GAACCAAGAGACGTCCT	4926				
QY	601	AAAGAAAAGTTT	GCAAGAGCAGCAAGAGATTTT	GAACCAAGAGAGGCTGATTC	CAAGAAA	660				
Db	4927	AAAGAAAAGTTT	GCAAGAGCAGCAAGAGATTTT	GAACCAAGAGAGGCTGATTC	CAAGAAA	4986				
QY	661	AATTTAAAGAAAAGAAA	GAACCAAGAGATTTT	TACAGAGATTTT	ATATGATTCGCTTTA	720				
Db	4987	AATTTAAAGAAAAGAAA	GAACCAAGAGATTTT	TACAGAGATTTT	ATATGATTCGCTTTA	5046				
QY	721	GAATATACAGCTT	ATAGAACTTCCATCAGAAAATG	GAACGATTTT	ATATACCAATCAA	780				
Db	5047	GAATATACAGCTT	ATAGAACTTCCATCAGAAAATG	GAACGATTTT	ATATACCAATCAA	5106				
QY	781	TCCTCTTTT	TAAGCTCAGAACACAGAGGATTT	GAAGATTTT	CAAGAAATTTCTATATA	840				
Db	5107	TCCTCTTTT	TAAGCTCAGAACACAGAGGATTT	GAAGATTTT	CAAGAAATTTCTATATA	5166				
QY	841	GAAGAAAACAAAT	TGAGAACTTATTA	CAACAAAGTTT	GAAGAGCAAGAGGATTT	900				
Db	5167	GAAGAAAACAAAT	TGAGAACTTATTA	CAACAAAGTTT	GAAGAGCAAGAGGATTT	5226				
QY	901	GGACATCTT	TGAGAAAGAAAGATGTTT	CAATATAAACCAAGAACAA	AAAGAGATTAATCT	960				
Db	5227	GGACATCTT	TGAGAAAGAAAGATGTTT	CAATATAAACCAAGAACAA	AAAGAGATTAATCT	5286				
Db	5287	GCTGACATATCAAAAT	CACTATTCATAGAGACAGTAAATTTT	CTGATGTTTAAATGATTTT	CAAA	5346				
QY	1021	ATATAGTAAGTATGAGATG	GAATGAATGAAGTGAATATG	CAATTCATTAATATGATGAAGAA	1080					
Db	5347	ATATAGTAAGTATGAGATG	GAATGAATGAAGTGAATATG	CAATTCATTAATATGATGAAGAA	5406					

QY	1081	GAAGATGATGAAGACTTGAACGAATTAAAGCCTATTTGCAATATGACAAATTTCCAAAT	1140
Db	5407	GAAGATGATGAAGACTTGAACGAATTAAAGCCTATTTGCAATATGACAAATTTCCAAAT	5466
QY	1141	GAAGAAAACATAGGAATTTATTAAGAACTAGAAAGTTTGAATAGAGAAAAAATGAATTTA	1200
Db	5467	GAAGAAAACATAGGAATTTATTAAGAACTAGAAAGTTTGAATAGAGAAAAAATGAATTTA	5526
QY	1201	GATGATTTAGATGAAGAAATGAAAAATCATCAGAAAGATTATCTGAAAGAAAAATATAAA	1260
Db	5527	GATGATTTAGATGAAGAAATGAAAAATCATCAGAAAGATTATCTGAAAGAAAAATATAAA	5586
QY	1261	AAAGGAAAGAAATATGAAAAAACAAGAGATATAATTTTAAACCAATGATPAAAGTTTG	1320
Db	5587	AAAGGAAAGAAATATGAAAAAACAAGAGATATAATTTTAAACCAATGATPAAAGTTTG	5646
QY	1321	TATGATGAGCATTTAAAAAATATATAAATGATAGAGAGTTTAAATAGGAAAAAGAAAAA	1380
Db	5647	TATGATGAGCATTTAAAAAATATATAAATGATAGAGAGTTTAAATAGGAAAAAGAAAAA	5706
QY	1381	TTCAATAAATCATTTTTCATATATTGACGAGACATGAAATTTTACAGATCGTGAT	1440
Db	5707	TTCAATAAATCATTTTTCATATATTGACGAGACATGAAATTTTACAGATCGTGAT	5766
QY	1441	GAGTTATCTGAAGATATACTAAATATTTTATGAAACTATAA	1482
Db	5767	GAGTTATCTGAAGATATACTAAATATTTTATGAAACTATAA	5808
RESULT 10	AE014834	253001 bp DNA linear INV 07-OCT-2002	
LOCUS	AE014834	Plasmodium falciparum 3D7 chromosome 10 section 6 of 7 of the	
DEFINITION	complete sequence.		
ACCESSION	AE014834	AE014834	
VERSION	AE014834.1	GI:23495173	
KEYWORDS			
SOURCE			
ORGANISM		Plasmodium falciparum 3D7 Plasmodium falciparum 3D7 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
REFERENCE	1	(bases 1 to 253001)	
AUTHORS	Gardner,M.J., Hall,N., Fung,E., White,O., Berrihan,M., Hyman,R.W., Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kye,S., Chan,M.-S., Nene,V., Shalom,S.J., Sub,B., Peterson,J., Anguilo,S., Percec,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Franholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummins,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.		
TITLE		Genome sequence of the human malaria parasite Plasmodium falciparum	
JOURNAL	Nature 419 (6906), 498-511 (2002)		
PUBMED	12368864		
REFERENCE	2	(bases 1 to 253001)	
AUTHORS	Gardner,M.J.		
TITLE	Direct Submision		
JOURNAL	Submitted (13-SEP-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA		
FEATURES			
source		1.. 253001 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="10" /note="AT rich" /rpt_type="random complement (113..136) /rpt_type=tandem complement (113..136) /rpt_unit="(a)n"	
repeat_region			
repeat_region			
repeat_region			

[illegible]

```
repeat_region 2808..2966 /rpt_type=tandem  
repeat_region 3142..3171 /rpt_type=tandem  
repeat_region /rpt_unit="(a)n"  
complement(3274)..3347 /note="AT_rich"  
repeat_region /rpt_type=tandem  
complement(3391)..3437 /rpt_unit="(ta)n"  
repeat_region complement(3439)..3503 /rpt_type=tandem  
repeat_region /rpt_unit="(a)n"  
3439..3500 /rpt_type=tandem  
repeat_region 3440..3500 /rpt_type=tandem  
complement(3551)..3683 /rpt_type=tandem  
repeat_region /rpt_unit="(tataa)n"  
complement(3710)..3745 /note="AT_rich"  
repeat_region /rpt_type=tandem  
complement(3763)..3824 /rpt_unit="(ta)n"  
3763..3811 /rpt_type=tandem  
complement(3871)..4025 /note="AT_rich"  
repeat_region /rpt_type=tandem  
complement(3922)..3987 /rpt_unit="(ca)n"  
3925..3987 /rpt_type=tandem  
complement(4037)..4082 /rpt_type=tandem  
repeat_region /rpt_unit="(tataa)n"  
4043..4082 /rpt_type=tandem  
4043..4080 /rpt_type=tandem  
complement(4084)..4104 /rpt_type=tandem  
complement(4135)..4181 /rpt_type=tandem  
repeat_region /rpt_unit="(ta)n"  
complement(4182)..4313 /note="AT_rich"  
repeat_region /rpt_type=tandem  
4364..4466 /rpt_type=tandem  
4372..4457 /rpt_type=tandem  
4377..4472 /rpt_type=tandem  
repeat_region /rpt_unit="(tataa)n"  
complement(4512)..4655 /note="AT_rich"  
repeat_region /rpt_type=tandem  
complement(4670)..4775 /rpt_unit="(ta)n"  
4675..4768 /rpt_type=tandem  
repeat_region 4870..4912 /rpt_type=tandem  
complement(4870)..4912 /rpt_type=tandem  
repeat_region /rpt_unit="(ca)n"
```

```
repeat_region complement(4918)..5000 /rpt_type=tandem  
repeat_region /rpt_unit="(tataa)n"  
gene /locus_tag="PF10_0308"  
mRNA join(<5162..5350,5488..5535,5617..5802,5902..5981,  
6099..6162,6254..6349,6463..6540,6727..>8973)  
CDS join(5162..5350,5488..5535,5617..5802,5902..5981,  
6099..6162,6254..6349,6463..6540,6727..8973)  
/locus_tag="PF10_0308"  
/product="hypothetical protein"  
Query Match 94.9%; Score 1406.8; DB 2; Length 253001;  
Best Local Similarity 96.8%; Pred. No. 4,2e-151;  
Matches 1435; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
1 CAAAGCAACAAAGCGATTGAGAACAAAGAGACGTGCTAAAGAAAAGTTGCAAGAACAA  
179846 CAGGGCCACAAAGCGATTGAGAACAAAGAGACTAGCTAAAGAAAAGTTACGGGCCAA 179905  
QY CAAAGCGATTGAGAACAAAGATAGACTTGCTAAAGAAAAGTTACAAAGCAGCAAGCGAT 120  
179906 CAAAGCATCTAGAACAAAGAGAGACTAGCTAAAGAAAAGTTACAGGGGCAACAAAGCGAT 179965  
Db 121 TTGAAACAAAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATCTTAGAACAA 180  
179966 TTAGAACAAAGAGACTTGCTAAAGAAAAGTTGCAAGAGGACAAAGCGATTTAGAACAA 180025  
QY GAGAGCGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCT 240  
180026 GAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCTA 180085  
QY GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATGCTGCTAAAGAA 300  
180086 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGCAGACTTGCTAAAGAA 180145  
QY 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGATGCTGCTAAAGAA 300  
180086 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGCAGACTTGCTAAAGAA 180145  
QY 301 AAGTTACAAAGAGCAGCAAGCGATTTAGAACAAAGAGAGCTGCTAAAGAAAAGTTGCCAA 360  
180146 AAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCTAGCTAAAGAAAAGTTACAG 180205  
QY 361 GAAACAAGAAAGCGATTTAGAACAAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAA 420  
180206 GGGCAGCATAGCGATTTAGAACAAAGCAGCATTTAAAGAACGTTTGCAAGAACAA 180265  
QY 421 AGCGATTTAGAACAAAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 480  
180266 AGCGATTTAGAACAAAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA 180325  
QY 481 GAAACAAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG 540  
180326 GAAACAAGAGAGCGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAG 180385  
QY 541 AGACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCT 600  
180386 AGACGTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGCGTCT 180445  
QY 601 AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCGTGAATACGAAAAAA 660  
180446 AAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCGTGAATACGAAAAAA 180505  
QY 661 AATTAGAAAAGAAAAGAAATGAGATATATTACAGAGATTTATATGTCGTTTA 720  
180506 AATTAGAAAAGAAAAGAAATGAGATATATTACAGAGATTTATATGTCGTTTA 180565  
QY 721 GAAATACAGCTATAGAACTTCCATCAGAAAAGAAAGTGGATTTATATACCATCAA 780  
180566 GAAATACAGCTATAGAACTTCCATCAGAAAAGAAAGTGGATTTATATACCATCAA 180625  
QY 781 TCTTCTTACCTCAGAGCAACAGAGGAAATGATAGATTTCAAGAAATATCTATATATA 840  
180626 TCTTCTTACCTCAGAGCAACAGAGGAAATGATAGATTTCAAGAAATATCTATATATA 180685
```

QY 841 GAAAAACAAATAGAGAAATCTATTACAACAAATGTTGAAAGAGAGGATATACATATAA 900
|
|
|
Db 180866 GAAAAATCAAAATAGAGAAATCTATTACAACAAATGTTGAAAGAGAGGATATATATAA 180745
|
|
|
QY 901 GGACATCTTTGAAAGAAAGAAAGATGTTCAATMAAACAGAAACAAAGAAAGATMAATCT 960
|
|
|
Db 180746 GGACATCTTTGAAAGAAAGAAAGATGTTCAATMAAACAGAAACAAAGAAAGATMAATCT 180805
|
|
|
QY 961 GCTGACATACAAATCTATCATTTAGAGACAGTMAATTTTCTGATGTTAAATGATTTTCAA 1020
|
|
|
Db 180806 GCTGACATACAAATCTATCATTTAGAGACAGTMAATTTTCTGATGTTAAATGATTTTCAA 180865
|
|
|
QY 1021 ATTAAGTATGATGAGATGATMAATTAAGTCTGAATTAAGACATTCATTAATAGATGAAGA 1080
|
|
|
Db 180866 ATTAAGTATGATGAGATGATMAATTAAGTCTGAATTAAGACATTCATTAATAGATGAAGA 180925
|
|
|
QY 1081 GAAGATGATGAAAGACTTGAACGAATTTAAGCCTATTGTGCAATATGACAAATTTCCAAAGAT 1140
|
|
|
Db 180926 GAAGATGATGAAAGACTTGAACGAATTTAAGCCTATTGTGCAATATGACAAATTTCCAAAGAT 180985
|
|
|
QY 1141 GAAGAAACATAGAAATTTATTAAGAACTAGAAATTTGATAGAGAAAAATGAATAATTA 1200
|
|
|
Db 180986 GAAGAAACATAGAAATTTATTAAGAACTAGAAATTTGATAGAGAAAAATGAATAATTA 181045
|
|
|
QY 1201 GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGATTAATCTGAAGAAAAATTAATA 1260
|
|
|
Db 181046 GATGATTTAGATGAAGAAATAGAAAAATCATCAGAAAGATTAATCTGAAGAAAAATTAATA 181105
|
|
|
QY 1261 AAAGAAAGAAATATGAAAAAAACAAAGATAATTTTAAACCAATGATMAAGTTTG 1320
|
|
|
Db 181106 AAAGAAAGAAATATGAAAAAAACAAAGATAATTTTAAACCAATGATMAAGTTTG 181165
|
|
|
QY 1321 TATGATGACATATTTAAAAATTTAAAAATGATTAAGCAGTTAAATGAAGAAAAAGAAAA 1380
|
|
|
Db 181166 TATGATGACATATTTAAAAATTTAAAAATGATTAAGCAGTTAAATGAAGAAAAAGAAAA 181225
|
|
|
QY 1381 TTCATTAATCATCTGTTGTTATATTTGACGAGACATGAATTTTACAGATCGTGAT 1440
|
|
|
Db 181226 TTCATTAATCATCTGTTGTTATATTTGACGAGACATGAATTTTACAGATCGTGAT 181285
|
|
|
QY 1441 GAGTATCTGAAGATATTAATTAATTTTATGAACATATA 1482
|
|
|
Db 181286 GAGTATCTGAAGATATTAATTAATTTTATGAACATATA 181327
|
|
|
RESULT 11
A28740 950 bp DNA linear PAT 04-JUN-1995
LOCUS A28740 CDNA for LSA-R-NR protein.
DEFINITION A28740
ACCESSION A28740
VERSION A28740.1 GI:1247512
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 950)
TITLE LIVER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING
EPIPTOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES
JOURNAL Patent: WO 9213884-A 29 20-AUG-1992;
FEATURES
Location/Qualifiers
1..950
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 64.1%; Score 950; DB 6; Length 950;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AAACGATCTAGAACAAAGAGAGACGTGCTMAAGAAAAATTGCCAAGAACAAAGCGATT 70
|
|
|
Db 1 AAACGATCTAGAACAAAGAGAGACGTGCTMAAGAAAAATTGCCAAGAACAAAGCGATT 60
|
|
|
QY 71 TAGAACAAAGATAGACTTCTTAAGAAAAATTACAAGACAGCAAAAGCGATTTAGAACAG 130
|
|
|
Db 61 TAGAACAAAGATAGACTTCTTAAGAAAAATTACAAGACAGCAAAAGCGATTTAGAACAG 120
|
|
|
QY 131 AGACATCTGCTMAAGAAAAATTGCCAAGAACAAAGCGATTTAGAACAGACAGCGTG 190
|
|
|
Db 121 AGACATCTGCTMAAGAAAAATTGCCAAGAACAAAGCGATTTAGAACAGACAGCGTG 180
|
|
|
QY 191 CTTAAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGCGATTTAGAACAA 250
|
|
|
Db 181 CTTAAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGCGATTTAGAACAA 240
|
|
|
QY 251 AGTTGCCAAGAACAAAGCGATTTAGAACAAAGTAGACTTGTCTMAAGAAAAATTACAAG 310
|
|
|
Db 241 AGTTGCCAAGAACAAAGCGATTTAGAACAAAGTAGACTTGTCTMAAGAAAAATTACAAG 300
|
|
|
QY 311 AGCAGCAAAAGGATTTAGAACAAAGAGACGTGCTMAAGAAAAATTGCCAAGAACAA 370
|
|
|
Db 301 AGCAGCAAAAGGATTTAGAACAAAGAGACGTGCTMAAGAAAAATTGCCAAGAACAA 360
|
|
|
QY 371 GCGATTTAGAACAAAGAGACGTGCTMAAGAAAAATTGCCAAGAACAAAGCGATTTAG 430
|
|
|
Db 361 GCGATTTAGAACAAAGAGACGTGCTMAAGAAAAATTGCCAAGAACAAAGCGATTTAG 420
|
|
|
QY 431 AACCAAGAGACCTGCTTAAGAAAAATTGCCAAGAACAAAGCGATTTAGAACAAAGAG 490
|
|
|
Db 421 AACCAAGAGACCTGCTTAAGAAAAATTGCCAAGAACAAAGCGATTTAGAACAAAGAG 480
|
|
|
QY 491 GACGCTTAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGTGCTA 550
|
|
|
Db 481 GACGCTTAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGTGCTA 540
|
|
|
QY 551 AAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGAGACGTGCTMAAGAAAAAGT 610
|
|
|
Db 541 AAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGAGACGTGCTMAAGAAAAAGT 600
|
|
|
QY 611 TGCAGAGACGACGAAGAGATTTAGAACAAAGAGACGTGATTCGAAAAAAATTTAGAAA 670
|
|
|
Db 601 TGCAGAGACGACGAAGAGATTTAGAACAAAGAGACGTGATTCGAAAAAAATTTAGAAA 660
|
|
|
QY 671 GAAAAAAGGAACATGAGATATATTAAGAGAGATTTATGATGCTTTAGAAAAATACAG 730
|
|
|
Db 661 GAAAAAAGGAACATGAGATATATTAAGAGAGATTTATGATGCTTTAGAAAAATACAG 720
|
|
|
QY 731 CTATAGAACTTCCATCAGAAAAATGAAACGTGATATTTATATACACATCAATCTCTTAC 790
|
|
|
Db 721 CTATAGAACTTCCATCAGAAAAATGAAACGTGATATTTATATACACATCAATCTCTTAC 780
|
|
|
QY 791 CTCAGGACACAGAGGAAATGTTAGAGATTCACAGGAATATCTATATAGAAAAAACAA 850
|
|
|
Db 841 ATAGAGAACTTATTAACAACAATGTTGAAGGCGAAGGATATATACATAAGGACATCTTG 910
|
|
|
QY 911 AAGAAAGAAAGATGTTCAATMAAACAGAACAAAGAAAGATMAATCT 960
|
|
|
Db 901 AAGAAAGAAAGATGTTCAATMAAACAGAACAAAGAAAGATMAATCT 950
|
|
|
RESULT 12
AR162961 950 bp DNA linear PAT 17-OCT-2001
LOCUS AR162961
DEFINITION Sequence 32 from patent US 6270771.
ACCESSION AR162961
VERSION AR162961.1 GI:16233425
KEYWORDS

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 950)
AUTHORS Guerin-Marchand, C. and Druilhe, P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 620771-A 32 07-AUG-2001;
FEATURES
 source
 1..950
 /organism="unknown"
 /mol_type="unassigned DNA"
ORIGIN
Query Match 64.1%; Score 950; DB 6; Length 950;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAAGCGATTCTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATT 70
DB 1 AAAGCGATTCTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATT 60
QY 71 TAGAACAAAGATAGACTTGTCTAAAGAAAAGTTACAGAGCAGCAAGCGATTAGAACAG 130
DB 61 TAGAACAAAGATAGACTTGTCTAAAGAAAAGTTACAGAGCAGCAAGCGATTAGAACAG 120
QY 131 AGAGACTTGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTG 190
DB 121 AGAGACTTGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTG 180
QY 191 CTAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAA 250
DB 181 CTAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAA 240
QY 251 AGTTGCAAGAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTAGAACAG 310
DB 241 AGTTGCAAGAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTAGAACAG 300
QY 311 AGCAGCAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAA 370
DB 301 AGCAGCAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAA 360
QY 371 GCCATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAG 430
DB 361 GCCATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAG 420
QY 431 AACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAG 490
DB 421 AACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAG 480
QY 491 GACGTGCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTGCTA 550
DB 481 GACGTGCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTGCTA 540
QY 551 AAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTGCTAAAGAAAAGT 610
DB 541 AAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTGCTAAAGAAAAGT 600
QY 611 TGCAAGAGCAGCAAGAGATTAGAACAAAGAGCGTGATACGAAAAAAAATTTAGAAA 670
DB 601 TGCAAGAGCAGCAAGAGATTAGAACAAAGAGCGTGATACGAAAAAAAATTTAGAAA 660
QY 671 GAAAAAGAGACATGAGATATATTAGCAGAGATTATATGTCGTTTAAAGAAATACAG 730
DB 661 GAAAAAGAGACATGAGATATATTAGCAGAGATTATATGTCGTTTAAAGAAATACAG 720
QY 731 CTATAGAACTTGCATCAGAAAAATGAACTGATATTATATACCATATATCTCTTTAC 790
DB 721 CTATAGAACTTGCATCAGAAAAATGAACTGATATTATATACCATATATCTCTTTAC 780
QY 791 CTCAGAGCAACAGAGGAAATAGTAGATTCCAGAGAAATATCTATAATAGAAAAACAA 850
DB 781 CTCAGAGCAACAGAGGAAATAGTAGATTCCAGAGAAATATCTATAATAGAAAAACAA 840

QY 851 ATAGAGATCTATTACAACAATGTTGAAGAGAGGATATATACATAAGAGCATCTTG 910
DB 841 ATAGAGATCTATTACAACAATGTTGAAGAGAGGATATATACATAAGAGCATCTTG 900
QY 911 AAGAAAAGAGATGTTCAATTAACAGAACAAAAAGAGATAATCT 960
DB 901 AAGAAAAGAGATGTTCAATTAACAGAACAAAAAGAGATAATCT 950
RESULT 13
AR178250 950 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 32 from patent US 6319502.
DEFINITION AR178250
ACCESSION AR178250 GI:20219388
VERSION AR178250.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 950)
AUTHORS Guerin-Marchand, C. and Druilhe, P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6319502-A 32 20-NOV-2001;
FEATURES
 source
 1..950
 /organism="unknown"
 /mol_type="unassigned DNA"
ORIGIN
Query Match 64.1%; Score 950; DB 6; Length 950;
Best Local Similarity 100.0%; Pred. No. 1.8e-98;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAAGCGATTCTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATT 70
DB 1 AAAGCGATTCTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATT 60
QY 71 TAGAACAAAGATAGACTTGTCTAAAGAAAAGTTACAGAGCAGCAAGCGATTAGAACAG 130
DB 61 TAGAACAAAGATAGACTTGTCTAAAGAAAAGTTACAGAGCAGCAAGCGATTAGAACAG 120
QY 131 AGAGACTTGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTG 190
DB 121 AGAGACTTGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTG 180
QY 191 CTAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAA 250
DB 181 CTAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAA 240
QY 251 AGTTGCAAGAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTAGAACAG 310
DB 241 AGTTGCAAGAACAAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTAGAACAG 300
QY 311 AGCAGCAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAA 370
DB 301 AGCAGCAAGCGATTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAA 360
QY 371 GCCATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAG 430
DB 361 GCCATTTAGAACAGAGAGCGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAG 420
QY 431 AACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAG 490
DB 421 AACAGAGAGACTTGTCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAG 480
QY 491 GACGTGCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTGCTA 550
DB 481 GACGTGCTAAAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTGCTA 540
QY 551 AAGAAAAGTTGCAGAGCAACAAAGCGATTAGAACAGAGAGCGTGCTAAAGAAAAGT 610

Db 541 AAGAAAGTTCGACAAACAAACGATTTAGAACAGAGACGCTGCTAAAGAAAGT 600
Qy 611 TGCAAGAGCAGCAAGAGATTTAGAACAAAGAGCGCTGATCGAAGAAAAATTTAGAAA 670
Db 601 TGCAAGAGCAGCAAGAGATTTAGAACAAAGAGCGCTGATCGAAGAAAAATTTAGAAA 660
Qy 671 GAAAAAGCAATGAGATATATATTTAGCAGAGATTTATATGTGCTTTAGAAATACGAG 730
Db 661 GAAAAAGCAATGAGATATATTTAGCAGAGATTTATATGTGCTTTAGAAATACGAG 720
Qy 721 CTATGAACTTCATCAGAAAATGAACGCTGATTTATATACCATCATCTTCTTTAC 790
Db 721 CTATGAACTTCATCAGAAAATGAACGCTGATTTATATACCATCATCTTCTTTAC 780
Qy 791 CTCAGAGACAGAGAGGATAGATTTCCAGGAAATATCTAATAGAAAAACAA 850
Db 781 CTCAGAGACAGAGAGGATAGATTTCCAGGAAATATCTAATAGAAAAACAA 840
Qy 851 ATAGAGAACTTATTAACAACAAATGTTGAAGACGAAAGGATATACATTAAGACATCTTG 910
Db 841 ATAGAGAACTTATTAACAACAAATGTTGAAGACGAAAGGATATACATTAAGACATCTTG 900
Qy 911 AAGAAAGCAAGATGCTTCAATPAAACCAACAAACAAAGACATTAATCT 960
Db 901 AAGAAAGCAAGATGCTTCAATPAAACCAACAAACAAAGACATTAATCT 950

RESULT 14
PFALSAAC 843 bp DNA linear INV 19-AUG-1995
LOCUS PNG1-28a, PNG1-28c, BRA1-19e liver stage-specific antigen 1
DEFINITION (LSA-1) gene, 3' end of cds.
ACCESSION L40885.1 GI:725278
VERSION 1.0
KEYWORDS exoerythrocytic; liver stage-specific antigen; liver stage-specific antigen 1; malaria.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
REFERENCE 1 (bases 1 to 843)
AUTHORS Yang, C., Shi, Y.P., Udhayakumar, V., Alpers, M.P., Pova, M.M., Hawley, W.A., Collins, M.E. and Lal, A.A.
TITLE Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates
JOURNAL Mol. Biochem. Parasitol. 71 (2), 291-294 (1995)
PUBMED 7477115
COMMENT Original source text: Plasmodium falciparum (individual isolate Kenyan) (clone: KEN1-56b) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG1-28a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG1-28c) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN1-21b) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KEN1-40h) blood stage parasite DNA; and Plasmodium falciparum (individual isolate Kenyan) (clone: BRA1-19e) blood stage parasite DNA.
FEATURES
source
1. 843
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/isolate="Kenyan"
/db_xref="taxon:5833"
/clone="KEN1-56b"
/clone="PNG1-28a"
/clone="PNG1-28c"
/clone="KEN1-21b"

variation
/gene="LSA-1"
/note="clone PNG1-28a"
/replace="c"
228
FMK"

ORIGIN
Query Match 56.9%; Score 843; DB 2; Length 843;
Best Local Similarity 100.0%; Pred. No. 2.3e-86;
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

gene
CDS
/clone="KEN1-40h"
/clone="BRA1-19e"
/dev_stage="blood stage parasite"
1. 843
/gene="LSA-1"
<1. 843
/gene="LSA-1"
/codon_start=1
/product="liver stage-specific antigen 1"
/protein_id="AAB59231.1"
/db_xref="GI:350664"
/translation="RKADTKNLRKKEHEDILADLYGRLEIPIELPSRENGCYI
PHQSLPDQNGNSRDSKEISIEKTNKESITTVNEGRBDLHKHLEKKQSGIKPEQ
KEDXSADIQNHTEITVNI SDVNDPOISYDEIEISAEYDLSIDEEDDELDIEFKPIV
QYDNFDEENIGIYKELEDLEKNEEDLDDEGIEKSESELSSEBKIKKQKYEKTDN
NFKPNDKSLYDEHIKKYKNDKQVKEKEKFIKSLFHPIDGNEILQIVDELSIDITKY
FMK"

Qy 640 AAGAAAGCTGATPACGAAAAAATTTAGAAAGAAAAAGAAACATGAGATATATTAGCA 699
Db 1 AAGAAAGCTGATPACGAAAAAATTTAGAAAGAAAAAGAAACATGAGATATATTAGCA 60
Qy 700 GAGGATTTATATGCTGCTTTAGAAATACCACTATAGAACTTCATCAGAAAAATGAACGT 759
Db 61 GAGGATTTATATGCTGCTTTAGAAATACCACTATAGAACTTCATCAGAAAAATGAACGT 120
Qy 760 GGATTTATATACCATCAATCTTCTTACTCAGACCAACAGGCGGAAATGTAGAGAT 819
Db 121 GGATTTATATACCATCAATCTTCTTACTCAGACCAACAGGCGGAAATGTAGAGAT 180
Qy 820 TTCAGGAAATATCTATATATGAAAAAACAATAGAGAACTATATACCAACAAATGTTGAA 879
Db 181 TTCAGGAAATATCTATATATGAAAAAACAATAGAGAACTATATACCAACAAATGTTGAA 240
Qy 880 GGACGAGAGATATCATATAAGACATCTTGAAGAAAAAGAAAGATGTTCAATAAAAACCA 939
Db 241 GGACGAGAGATATCATATAAGACATCTTGAAGAAAAAGAAAGATGTTCAATAAAAACCA 300
Qy 940 GAAACAAAAAGATTAATCTGCTGACATACAAAATCTACTTGAACAGTAATATT 999
Db 301 GAAACAAAAAGATTAATCTGCTGACATACAAAATCTACTTGAACAGTAATATT 360
Qy 1000 TCTGATGTTAATGATTTCAATTAAGTAAGTGAAGTGAATGAATGCTGAATATGAC 1059
Db 361 TCTGATGTTAATGATTTCAATTAAGTAAGTGAAGTGAATGAATGCTGAATATGAC 420
Qy 1060 GATTCAATTAATGATGAGAGAGAGATGATGAGACTTAGAGCAATTTAAGCCTATTGTC 1119
Db 421 GATTCAATTAATGATGAGAGAGAGATGATGAGACTTAGAGCAATTTAAGCCTATTGTC 480
Qy 1120 CAATATGACAAATTTCCAGAGATGAGAAAAATGAGAAATTTATTAAGAACTAGAAAGTTTG 1179
Db 481 CAATATGACAAATTTCCAGAGATGAGAAAAATGAGAAATTTATTAAGAACTAGAAAGTTTG 540
Qy 1180 ATAGAGAAAAATGAAAAATTTAGATGATTTAGTGAAGGAAATGAAGAAATCATCAGAGAA 1239
Db 541 ATAGAGAAAAATGAAAAATTTAGATGATTTAGTGAAGGAAATGAAGAAATCATCAGAGAA 600
Qy 1240 TTATCTGAGAAAAAATATATATATGAGAAAGAAATATGAGAAAGCAAGATATATTTT 1299
Db 601 TTATCTGAGAAAAAATATATATATGAGAAAGAAATATGAGAAAGCAAGATATATTTT 660
Qy 1300 AAACCAAAATGATTAAGTTTGTATGATGAGCATATTTAAAAATATTAATATTAAGCAG 1359

Db 661 AACCAATGATTAAGTTGTATGATGACATATTAATAATATTAATAATGATACAG 720

Qy 1360 GTTAATAAGAAAAAGAAAAATTCATAAATCATTTGTTTCATATATTGACGAGACAT 1419

Db 721 GTTAATAAGAAAAAGAAAAATTCATAAATCATTTGTTTCATATATTGACGAGACAT 780

Qy 1420 GAAATTTTACGATCGTGAGACTTATCTGAATATTAATAATTTTATGAACTA 1479

Db 781 GAAATTTTACGATCGTGAGACTTATCTGAATATTAATAATTTTATGAACTA 840

Qy 1480 TAA 1482

Db 841 TAA 843

RESULT 15

PFAL5AB 843 bp DNA linear INV 19-AUG-1995

LOCUS

DEFINITION Plasmodium falciparum (clones KENO-11a, -11b, -37d, -39a, -44e, -51a, -53c, -53d, -56a, PNG0-5h, -7a, -8f, -19a, -26b, -86h, -86i, BRA0-5a) liver stage-specific antigen 1 (LSA-1) gene, 3' end of cds.

ACCESSION L40884.1 GI:725277

VERSION

KEYWORDS exerythrocytic; liver stage-specific antigen; liver stage-specific antigen 1; malaria.

SOURCE Plasmodium falciparum (malaria parasite P. falciparum)

ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 843) Yang, C., Shi, Y. P., Udhayakumar, V., Alpers, M. P., Povoa, M. M., Hawley, W. A., Collins, W. E. and Lal, A. A. Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates

AUTHORS

TITLE

JOURNAL

COMMENT

Original source text: Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-44e) DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-51a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-7a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-5h) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-11a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-11b) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-37d) DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-39a) DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-8f) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-19a) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-26b) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-86h) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: PNG0-86i) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-53c) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-53d) blood stage parasite DNA; Plasmodium falciparum (individual isolate Kenyan) (clone: KENO-56a) blood stage parasite DNA; and Plasmodium falciparum (individual isolate Kenyan) (clone: BRA0-5a) blood stage parasite DNA.

Clone name abbreviations: KEN for Kenya; PNG for Papua New Guinea and BRA for Brazil. For each clone, the number following the '-', PNG or BRA represents group number, the number following the '-', denotes the patient's blood sample number and the lower case letter represents recombinant clones from the corresponding isolates.

FEATURES

Source

1. .843

/organism="Plasmodium falciparum"

/mol_type="genomic DNA"

/isolate="Kenyan"

/db_xref="taxon:5833"

/clone="KENO-51a"

/clone="KENO-44e"

/clone="KENO-7a"

/clone="PNG0-7a"

/clone="PNG0-5h"

/clone="KENO-11a"

/clone="KENO-11b"

/clone="KENO-37d"

/clone="KENO-39a"

/clone="PNG0-8f"

/clone="PNG0-19a"

/clone="PNG0-26b"

/clone="PNG0-86h"

/clone="PNG0-86i"

/clone="KENO-53c"

/clone="KENO-53d"

/clone="KENO-56a"

/clone="BRA0-5a"

/dev_stage="blood stage parasite"

1. .843

/gene="LSA-1"

/db_xref="GI:725277"

/gene="LSA-1"

/codon_start=1

/product="liver stage-specific antigen 1"

/protein_id="AAB59230.1"

/db_xref="GI:950863"

/translation="RKADYTKNLERKKEHGVLAIEDYGRLEIPALEPSENERGYI PHOSLIPODNRGSRDSKEISIEKTNRESITTNVGRDRIHGHLEEKDGSIKPEQ KEDSLDIQHTLETYNI SPVNDFOISKEYDEISAEYDSDLIBEDBDLDFEKPV QYNPDENIGIKYKELEDIEKNENDLDDEGIEKSESESESEKIKKGRYEKTQN NPKPNDLSLYDEHIKKYKNDKQVKEKEKIKSLFHPFDQDNEILQVLSLEDITKY FMKL"

variation

60

/gene="LSA-1"

/note="clone BRA0-5a"

/replace="g"

ORIGIN

Query Match 56.8%; Score 841.4; DB 2; Length 843;

Best Local Similarity 99.9%; Pred. No. 3.5e-86;

Matches 842; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 640 AGCAAGCTGATACGAAAAAATTTAGAAAAAGAAAAAGAACCTGAGATATATTGCA 699

Db 1 AGCAAGCTGATACGAAAAAATTTAGAAAAAGAAAAAGAACCTGAGATATATTGCA 60

Qy 700 GAGGATTTATATGTCGTTTAGAAATACCACTATAGAACTTCATCAGAAAAATGAACT 759

Db 61 GAGGATTTATATGTCGTTTAGAAATACCACTATAGAACTTCATCAGAAAAATGAACT 120

Qy 760 GATATTTATATACCATCAATCTCTTACCTCAGACAACAGAGGAAATAGTAGAT 819

Db 121 GATATTTATATACCATCAATCTCTTACCTCAGACAACAGAGGAAATAGTAGAT 180

Qy 820 TCACAGAAATATTTATTAATGAAAAACAATGAGAACTTATTAACAATAATGTTGAA 879

Db 181 TCACAGAAATATTTATTAATGAAAAACAATGAGAACTTATTAACAATAATGTTGAA 240

Qy 880 GGAGGAAGGATATACATAAAGACATCTTGAAGAAAAAGAAAGATGGTTCAATTAACA 939

Db 241 GGAGGAAGGATATACATAAAGACATCTTGAAGAAAAAGAAAGATGGTTCAATTAACA 300

Qy 940 GAACAAAAAGATAATCTGCTGACATCAAAATCATACATTAGAGACGTAATATT 999

Db 301 GAACAAAAAGATAATCTGCTGACATCAAAATCATACATTAGAGACGTAATATT 360

Qy 1000 TCTGATGTTAATGATTTTCAATTAAGTAAGTAGAGATGAATTAATGCTGGAATAGAC 1059

Db 361 TCTGATGTTAATGATTTTCAATTAAGTAAGTAGAGATGAATTAATGCTGGAATAGAC 420

Qy 1060 GATTCATTAATAGATGAAGAAAGATGATGAACCTTAGACGATTTAAGCCTATTGTG 1119

```

Db      421 GATTGATTAATGATGAGAGAGAGATGATGAAGACTTAGACGAATTTAAGCCATATGTG 480
QY      1120 CAATGTGACAAATTTCCAGATGAGAAAAACATAGCAATTTATTAAGAACTAGAAATTTG 1179
Db      481 CAATGTGACAAATTTCCAGATGAGAAAAACATAGCAATTTATTAAGAACTAGAAATTTG 540
QY      1180 ATAGAGAAAAATGAAAAATTTAGATGATTTAGATGAGAAATGAAAAATCATCAGAGAA 1239
Db      541 ATAGAGAAAAATGAAAAATTTAGATGATTTAGATGAGAAATGAAAAATCATCAGAGAA 600
QY      1240 TTATCTGAGAAAAATTAAGAAAAAGAAAAATATGAAAAACAAAGATATATATTTT 1299
Db      601 TTATCTGAGAAAAATTAAGAAAAAGAAAAATATGAAAAACAAAGATATATATTTT 660
QY      1300 AAACCAATGATTAAGATTTGATGATGACATATTAATAAATATTAATAAGTAAAGCAG 1359
Db      661 AAACCAATGATTAAGATTTGATGATGACATATTAATAAATATTAATAAGTAAAGCAG 720
QY      1360 GTTATTAAGAAAAAGAAAAATTCATTAATCATTTGTTCAATATATTGACGAGACAAT 1419
Db      721 GTTATTAAGAAAAAGAAAAATTCATTAATCATTTGTTCAATATATTGACGAGACAAT 780
QY      1420 GAAATTTTACAGATCGTGATGAGTTATCTGAAGATATTAATAAATTTTATGAACCTA 1479
Db      781 GAAATTTTACAGATCGTGATGAGTTATCTGAAGATATTAATAAATTTTATGAACCTA 840
QY      1480 TAA 1482
Db      841 TAA 843

```

Search completed: December 7, 2005, 01:48:02
 Job time : 7831 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_rnp model

Run on: December 7, 2005, 06:11:21, Search time 33 Seconds

(Without alignments)
8642.015 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 2486
Sequence: 1 CAAGACACACAAAGCGATCT.....AATATTTTATGAAACTATTA 1482

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 28346 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+np.model -DEV=xjh
-Q/cgn2.1/USPTO.spool.h/US09837344/runac.05122005.133407.29473/app_query.fasta.1.1671
-DB=PIR -OPMT=fastag -SUFFIX=rcp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -STRAT=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALLIGN=200 -THR_SCORE=spc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344 @CGN 1 1 40 @runac.05122005.133407.29473 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTR -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2406.5	96.8	1909	2	A45592 liver stage antigen
2	380.5	15.3	1407	1	A52858 trichohyalin - rab
3	380.5	15.3	1549	1	A40691 trichohyalin - she
4	376	15.1	1898	1	A45973 trichohyalin - hum
5	356	14.3	1017	2	T15598 hypothetical prote
6	353.5	14.2	737	2	T15597 hypothetical prote
7	332.5	13.4	1192	2	A71623 probable secreted
8	326.5	13.1	1979	1	S03166 myosin heavy chain
9	323.5	13.0	1039	2	S18199 myosin heavy chain
10	321	12.9	771	1	A34330 h-caldesmon - chic
11	319	12.8	1938	2	JCS421 smooch muscle myos
12	319	12.8	1972	2	JCS420 smooch muscle myos
13	318.5	12.8	1961	1	A61231 myosin heavy chain
14	317.5	12.8	1166	2	H86341 hypothetical prote

15	317.5	12.8	1937	2	I38055 myosin heavy chain
16	317.5	12.8	1992	2	A47297 myosin heavy chain
17	317	12.8	1432	2	B85431 trichohyalin like
18	316	12.7	1940	2	A29320 myosin heavy chain
19	314	12.6	1938	2	A59293 skeletal myosin he
20	314	12.6	1976	2	A59252 myosin heavy chain
21	313.5	12.6	1738	2	T14867 interaptin - blime
22	312.5	12.6	1959	2	A33977 myosin heavy chain
23	312	12.6	2007	1	B43402 myosin heavy chain
24	311.5	12.5	2139	2	T18296 myosin heavy chain
25	310.5	12.5	665	2	B71609 hypothetical prote
26	310.5	12.5	678	2	A54514 glutamic acid-rich
27	310	12.5	1963	1	MMKW myosin heavy chain
28	310	12.5	2116	2	A26655 myosin heavy chain
29	309.5	12.4	1956	2	T16416 hypothetical prote
30	309	12.4	1931	2	A59234 slow myosin heavy
31	309	12.4	1972	1	A41604 myosin heavy chain
32	307.5	12.4	1701	2	T09127 probable erythrocy
33	307	12.3	746	2	T47237 myosin II heavy ch
34	306.5	12.3	522	2	C96608 hypothetical prote
35	306.5	12.3	1790	2	S67593 transport protein
36	306	12.3	1999	1	S21801 myosin heavy chain
37	305.5	12.3	2017	1	A36014 myosin heavy chain
38	305.5	12.3	2057	2	S61477 nonmuscle myosin I
39	305	12.3	1964	2	A59282 myosin alpha heavy
40	304.5	12.2	1939	1	A46762 hypothetical prote
41	304	12.2	451	2	G70241 alpha cardiac myos
42	303.5	12.2	1938	2	T49464 myosin heavy chain
43	303.5	12.2	1939	2	T48175 mature-parasite-in
44	303	12.2	1526	2	A45605 hypothetical prote
45	302.5	12.2	1130	2	T34081

ALIGNMENTS

RESULT 1
A45592
liver stage antigen LSA-1 - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 22-Nov-1993 #sequence revision 02-Dec-1994 #text change 09-Jul-2004
C/Accession: S24597; A45592; S29393; S34842; B45592; D45592
R/Zhu, J.; Hollingdale, M.
Submitted to the EMBL Data Library, November 1990
A/Reference number: S24597
A/Accession: S24597
A/Molecule type: DNA
A/Residues: 1-1909 <ZHU>
A/Cross-references: UNIPROT:Q25893; UNIPARC:UPI0000081753; EMBL:X56203; NID:g9915; PID:
R/Zhu, J.; Hollingdale, M.R.
Mol. Biochem. Parasitol. 48, 223-226, 1991
A/Title: Structure of Plasmodium falciparum liver stage antigen-1.
A/Reference number: A45592; MUID:92107224; PMID:1840628
A/Accession: A45592
A/Molecule type: DNA
A/Residues: 1-195,638-668;1165-1215;1590-1909 <ZH2>
A/Cross-references: UNIPARC:UPI0000177376; UNIPARC:UPI0000177379; UNIPARC:UPI000017737A
A/Note: Sequence extracted from NCBI backbone (NCBIN:75010, NCBIN:75012, NCBIN:75014, N
R/Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Pataraopitkul, J.; Beaudoin
Nature 329, 164-167, 1987
A/Title: A liver-stage-specific antigen of Plasmodium falciparum characterized by gene
A/Reference number: S29393; MUID:87315391; PMID:3306406
A/Accession: S29393
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 323-387 <GUB1>
A/Cross-references: UNIPARC:UPI000017737C; EMBL:M28266
R/Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Pataraopitkul, J.; Beaudoin
Submitted to the EMBL Data Library, April 1992
A/Description: a liver-stage-specific antigen of Plasmodium falciparum characterized by
A/Reference number: S34842
A/Accession: S34842
A/Status: preliminary
A/Molecule type: DNA

A:Residues: 323-381, 'HKAI' <GUE2>
A:Cross-references: UNIPARC:UPI000017737D; EMBL:M28266
A:Note: difference at carboxyl end due to frameshift error
C:Comment: This protein is found as flocculent material in the parasitophorous vacuole.
C:Superfamily: trichohyalin; calmodulin repeat homology
C:Keywords: EF hand
F:154-1629/Region: 17-residue repeats (A-K-E-K-L-Q-E-Q-S-D-L-E-Q-E-R-R)

Alignment Scores:

Pred. No.:	2, 57e-119	Length:	1909
Score:	2406.50	Matches:	487
Percent Similarity:	89.89%	Conservative:	2
Best Local Similarity:	89.52%	Mismatches:	4
Query Match:	96.80%	Indels:	51
DB:	2	Gaps:	2

US-09-837-344-41 (1-1482) x A45592 (1-1909)

```

QY      1  CAAGAACAAACAAAGCGATTAGAACAAAGAGACGTGCTTAAAGAAAAGTTGCAGAACAA
DB      1366  GInGInGInGInSerAspLeuGInGInuArGArGAlaLySGInuLySGInGInGIn
QY      61  CAAAGCGATTTCAGAACAAATGACTT-----
DB      1386  GInSerAspLeuGInGInuArGArGAlaLySGInuLySGInGInGInGInuArGAr
QY      88  -----GCTAAAGAAAAGTTACAAAGACGACAAAGCGATTAGAACAA
DB      1406  LeuGInGInGInuArGArGAlaLySGInuLySGInGInGInGInSerAspLeuGIn
QY      130  GAGAGACTTGCTTAAAGAAAAGTTGCAGAACAAAGCGATTAGAACAAAGACGT
DB      1426  GInuArGArGAlaLySGInuLySGInGInGInGInSerAspLeuGInGInuArGAr
QY      190  GCTAAAGAAAAGTTGCAGAACAAACAAAGCGATTTCAGAACAAAGACGTCTTAAAGAA
DB      1446  AlAluLySGInuLySGInGInGInGInSerAspLeuGInGInGInuArGArGAlaLySGIn
QY      250  AAGTTGCAGAACAAACAAAGCGATTTCAGAACAAAGATGACTTGCTTAAAGAAAAGTTACAA
DB      1466  LySGInGInGInGInGInSerAspLeuGInGInGInuArGArGAlaLySGInuLySGIn
QY      310  GAGCAGCAAAAGCGATTTCAGAACAAAGACGTGCTTAAAGAAAAGTTGCAGAACAAACAA
DB      1486  GInGInGInuArGArGAspLeuGInGInuArGArGAlaLySGInuLySGInGInGIn
QY      370  AGCGATTTCAGAACAAAGACGTGCTTAAAGAAAAGTTGCAGAACAAACAAAGCGATTTC
DB      1506  SerAspLeuGInGInGInuArGArGAlaLySGInuLySGInGInGInGInSerAspLeu
QY      430  GAACAAAGAGAGACTTGCTTAAAGAAAAGTTGCAGAACAAACAAAGCGATTTCAGAACAG
DB      1526  GInGInGInuArGArGAlaLySGInuLySGInGInGInGInuArGArGAspLeuGInGIn
QY      490  AGACGTGCTTAAAGAAAAGTTGCAGAACAAACAAAGCGATTTCAGAACAGACGTGCT
DB      1546  ArGArGAlaLySGInuLySGInGInGInGInSerAspLeuGInGInuArGArGAla
QY      550  AAAAGAAAAGTTGCAGAACAAACAAAGCGATTTCAGAACAAAGACGT-----
DB      1566  LySGInuLySGInGInGInGInGInSerAspLeuGInGInGInuArGArGAlaLySGInu
QY      597  -----
DB      1586  LeuGInGInGInGInSerAspLeuGInGInGInuArGArGAlaLySGInuLySGInGIn
QY      598  -----GCTAAAGAAAAGTTGCAGAACAGACGACAAAGAA
DB      1606  GInGInuArGArGAspLeuGInGInGInuArGArGAlaLySGInuLySGInGInGInGIn
QY      628  GATTTCAGAACAAAGACGTGCTTAAAGAAAAGTTTCAGAACAAAGAAAGACATGGA
DB      1626  AspLeuGInGInuArGArGAlaLySGInuLySGInGInGInGInGInGInGInGInGIn

```

```

QY      688  GATATATTAGAGAGAGATTATATGTCGTTAGAAAATACAGACTTATGAACTTCATCA
DB      1646  AspValLeuAlaGInuAspLeuTyGlyArGInGInuLeuProAlaLeuLeuProSer
QY      748  GAAATGACGTCGATATTATATACACATCAATTTCTTACCTCAGACAAACAGAGG
DB      1666  GInuGInuArGArGAlaLySGInuLySGInGInGInSerAspLeuProGInAspAsnArgGly
QY      808  AATAGTAGAGATTCACAGAAATATCTTAATAGAAAACAAATAGAGAACTCTTCA
DB      1686  AsnSerAspAspSerLySGInuLySGInGInGInGInGInGInGInGInGInGInGIn
QY      868  ACAATGTTGAAGACGAAAGGATATATACATTAAGACACTCTTGAAGAAAAGAAAGATG
DB      1706  ThrAsnValGInGInuArGArGAspLeuLeuLySGInuLySGInGInuLySGInuLySGIn
QY      928  TCAATTAACCAAGAACAAAGAAAGATTAATCTGCTGACATPACAAATCATTCATTAG
DB      1726  SerLeuLySGInGInGInGInuAspLySGInGInGInGInGInGInGInGInGInGIn
QY      988  ACAGTAATATTTTCGATGTTAAAGATTTTCAATTAAGTAAGTAAGTAAGTAAGTAAGT
DB      1746  ThrValAsnLeuSerAspValAsnAspPheGInuLySGInGInGInGInGInGInGIn
QY      1048  GCTGATATGACGATTCATTAATAGATGAAGAAAGATGATGAACTTACGCAATTT
DB      1766  AlaGInuTyArAspAspSerLeuLeuAspGInuGInuAspAspGInuAspLeuAspGInu
QY      1108  AAGCTATTGTCATATATGACAAATTTCCAGATGAGAAAACATPAGAAATTTATTAAGAA
DB      1786  LysProLeuValGInuTyArAspAsnPheGInuAspGInuGInuLeuGlyLeuTySGIn
QY      1168  CTAGAAGATTGATAGAAAATGAAAATTTAGTGATTTAGAGAAAGTAAGAAA
DB      1806  LeuGInuAspLeuLeuGInuAsnGInuAsnLeuAspAspLeuAspGInuLySGInuLySGIn
QY      1228  TCATCAGAGAAATTTATCTGAGAAAATTAATAAGAAAAGAAAGAAATATGAAAAG
DB      1826  SerSerGInuGInuLeuSerGInuGInuLeuLySGInuLySGInuTySGInuLySGIn
QY      1288  GATAATATTTTAAACCAATGATTAAGTTGTATGATGACATATTTAAATAATATAA
DB      1846  AsnAsnAsnPheLySGInuAsnAspSerLeuTyArAspGInuLeuLeuLySGIn
QY      1348  AATGATTAAGCGCTTAAATTAAGAAAAGAAATTTCAATTAATCTTTCATATATT
DB      1866  AsnAspLySGInuValAsnLySGInuLySGInuPheLeuLySGInuLeuPheLeu
QY      1408  GACGAGACAATGAAATTTTACAGATCGTGAATGATGATGATGATGATGATGATGAT
DB      1886  AspGInuAspAsnGInuLeuGInuLeuValAspGInuLeuSerGInuAspLeuThrLySGIn
QY      1468  TTTATGAAACTA 1479
DB      1906  PheMetLySGInu 1909

```

RESULT 2

S28589
trichohyalin - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 12-Mar-1993 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
R:Fietz, M.J.; Rogers, G.B.
submitted to the EMBL Data Library, December 1992
A:Description: Examination of the gene encoding rabbit trichohyalin.
A:Reference number: S28589
A:Accession: S28589
A:Molecule type: DNA
A:Residues: 1-1407 <FIE>
A:Cross-references: UNIPROT:P37709; UNIPARC:UPI000013738C; EMBL:Z19092; NID:G1746; PIRN:
C:Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root st
Covalent modifications to this protein include conversion of arginine to citrulline and

1318 TGGATGATGAGCATATTATTAATAATATTAATGACGCTTAATTAAGCAAAAGCA 1377
Db 773 GlnTyrLeuGlnIuValGlnIuLeuGlnIuArgGlnIuGlnIuLeuGlnIuValArg 792
Qy 1378 AAA 1380
Db 793 Arg 793
RESULT 4
A45973
 trichohyalin - human
 C:Species: Homo sapiens (man)
 C:Date: 03-May-1994 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
 C:Accession: A45973
 R:Lee, S.C.; Kim, I.G.; Marekov, L.N.; O'Keefe, E.J.; Parry, D.A.D.; Steinert, P.M.
 J. Biol. Chem. 268, 12164-12176, 1993
 A:Title: The structure of human trichohyalin. Potential multiple roles as a functional E
 ed (cross-linking) protein.
 A:Reference number: A45973; MUID:93280194; PMID:7685034
 A:Accession: A45973
 A:Molecule type: DNA
 A:Residues: 1-1898 <LE>
 A:Cross-references: UNIPROT:007283; UNIPARC:UPI000013738B; GB:L09190; NID:G292835; PIDN
 A:Note: authors translated the codon AGG for residue 1714 as Pro
 C:Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root sh
 covalent modifications to this protein include conversion of arginine to citrulline and
 C:Genetics:
 A:Gene: GDB:7TH
 A:Cross-references: GDB:136223; OMIM:190370
 A:Map position: 1q21-1q21
 C:Superfamily: trichohyalin; calmodulin repeat homology
 C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <BF2>
Alignment Scores:
 Pred. No.: 5.43e-13 Length: 1898
 Score: 376.00 Matches: 137
 Percent Similarity: 47.39% Conservative: 117
 Best Local Similarity: 25.56% Mismatches: 174
 Query Match: 15.12% Indels: 108
 DB: 1 Gaps: 24
 US-09-837-344-41 (1-1482) x A45973 (1-1898)
Qy 1 CAAGAACAAACAAGCGATCTAGAACCAAGAGACGCTGCTAAAGAAAGTTGCAAGAACAA 60
Db 522 ArgIuGlnIuGlnIuLeuArgIuArgIuGlnIuArgIuArgIuGlnIuArgIuValArgGln 541
Qy 61 CAAGCGATTTAGAACCAAGATAGACTTCTTAAAGAAAGTTTAAAGACGACCAAGCGAT 120
Db 542 Gln-----GlnIuGlnIuArgIuLeuGlnIuArgIuValArgIuGlnIuLeuArg 558
Qy 121 TTAGAACAA--GAGAGACTTGCTTAAAGAAAGTTGCAAGAACCAAGCGATCTAGAA 177
Db 559 ArgIuGlnIuGlnIuArgIuLeuGlnIuLeuLeuValArgIuGlnIuValArgIuLeu 578
Qy 178 CAAGAGACGCTGCTTAAAGAAAGTTGCAAGAACCAAGC----- 219
Db 579 GlnIuValArgIuGlnIuArgIuLeuValArgIuGlnIuGlnIuValArgIuLeuValArg 598
Qy 220 GATTTAGAACCAAGAGACGCTGCTTAAAGAAAGTTGCAAGAACCAAGCGATTTAGAA 279
Db 599 LysValArgIuGlnIuArgIuArgIuGlnIuArgIuLeuValArgIuGlnIuGlnIuValArg 618
Qy 280 CAAGTTAGACTTGCTTAAAGAAAGTTTAAAGACGACGAAACGATTTAGACCAAGAGAGA 339
Db 619 Gln--ArgLeuValArgIuGlnIuValArgIuValArgIuValArgIuValArgIuValArg 633
Qy 340 CGTGCTTAAAGAAAGTTG-----CAAGAACCAAGCGATTTA--- 378
Db 634 ArgValArgIuValArgIuLeuValArgIuGlnIuProGlnIuGlnIuValArgIuGlnIuLeu 653
Qy 379 -----GAACCAAGAGACGCTGCTTAAAGAAAGTTGCAAGAACCAAGCGATTTA 429

Db 654 LysSerGluGluGlnGluGluLysArgArgHisGluGlnLeuArgArgGluGlnGlnLysArg 673
 |||||
 Oy 430 GAACAGAGACCTTGCTAAAGAAAAGTTCGACAAACA-----CAAAAC 474
 : : : : :
 Db 674 ArgGluGlnArgLeuLysArgGluGluGlnGluLysArgLeuGlnArgLeuLysArg 693
 : : : : :
 Oy 475 GATTTAACAACAAGACACCGCTCAAAAGAAAAGTTCGAAGAACAAACAAC----- 525
 : : : : :
 Db 694 GluHisGluGlnGluLysArgArgGluGlnGlnLeuAlaGluGlnGlnGlnAlaArg 713
 : : : : :
 Oy 526 -----GATTAGAACAAAGAG----- 540
 |||||
 Db 714 GluArgIleLysSerArgIleProLysTrpGlnTrpGlnLeuGlnSerGluAlaAspAla 733
 : : : : :
 Oy 541 AGACGTGCTAAAGAAAAGTTCGACAAACAAAC-----GAT 579
 : : : : :
 Db 734 ArgGlnSerLysValLeuLeuGluAlaProGlnAlaGlyArgAlaGluAlaProGlnGlu 753
 : : : : :
 Oy 580 TTAGAACCAAGAGACCGCTCAAAAGAAAAGTTCGA----- 615
 : : : : :
 Db 754 GlnGluGlnLysLysArgArgArgGlnSerGlnLeuGlnTrpGlnGlnGluLysArgAlaHis 773
 : : : : :
 Oy 616 GAGCAGCAAAAGATTATAGAACCAAGAGAGCGCTATACG---AAAAAAATTTTGAAGA 672
 |||||
 Db 774 ArgGlnGlnGlnGlnGlnGlnGlnArgGlnArgSerPheHisTrpGlnTrpGlnAlaGluGlu 793
 : : : : :
 Oy 673 AAAAAAGAACATGAGATATATATAGACAGAGATTATATAGTCGT----- 717
 |||||
 Db 794 LysSerGluLysGly-----ArgGlnArgLeuSerAlaArgProLeuLysGlu 810
 : : : : :
 Oy 718 -----TTAGAAATATCCAGCTATATAGACTTCCTCATCAGAAAATGAACCGTGATATTAT 768
 : : : : :
 Db 811 GlnArgGluLysArgGlnLeuArgAlaGluGlnArgGlnGlnArgGlnGlnArg-----Phe 828
 : : : : :
 Oy 769 ATACACATCATCATCTTCTTACCTCAGAGAACAGAGGAGATATAGAGATTCCAGGAA 828
 : : : : :
 Db 829 LeuProGlnGlnGlnGlnGlnGlnGlnArgGlnArgGlnArgGlnArgGlnArgGluLysGlu 848
 : : : : :
 Oy 829 ATATCTATATAGAAAAAACAAATAGA-----GAATCTATTACA 867
 : : : : :
 Db 849 LeuGlnPheLeuGlnGlnGlnGlnGlnGlnLeuGlnArgArgLysArgAlaGlnGlnLeuGln 868
 : : : : :
 Oy 868 ACAAATGTTGACAGA---CGAAGGATATACATATAAGACACTCTTGAAGAAAAGAAAGAT 924
 : : : : :
 Db 869 GluGlnGluLysArgGlyLeuGlnGlnLysArgGlnLysArgArgGlnGlnGlnArgGlnArgSerP 888
 : : : : :
 Oy 925 GGTTCATATAAACACAGAACAAAGAAAGATATAATCTGCTGACATACAAATATCATATACATA 984
 : : : : :
 Db 889 GlnLysTrpArgTrpGlnLeuGlnGlnGlnLysArg-----ArgArgHisThrLeu 906
 : : : : :
 Oy 985 GAGACAGTAAATATTTCTGATGTAAATGATTTCAAAATACGTAAGTATGAGATGAAATA 1044
 : : : : :
 Db 907 -----TyrLalysProAlaLeuGlnGlnLeu 916
 : : : : :
 Oy 1045 AGCTGCAATATGACGATTCTATTAATAGATCAAGAGACATGATGAAGACTTATGACGAA 1104
 : : : : :
 Db 917 ArgLysGlu-----GlnGlnLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnArgGlnGln 934
 : : : : :
 Oy 1105 TTTAAGCCTATTTGCAATATAGACATTTCCAAATGAGAGAAAACATAGCAATTTATATAA 1164
 : : : : :
 Db 935 -----ArgGlnLysArgArgArgGlnGlnGlnGlnGlnGlnLysArg 948
 : : : : :
 Oy 1165 GAACATGAAAGATTGATAGAGAAAAATCAAAAAATTTAGATGATTTAGTGAAGAAATAGAA 1224
 : : : : :
 Db 949 GluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 966
 : : : : :
 Oy 1225 AAATCATCAGAGAAATTATCTGAAGAAAAATATAAAAAAGGAAAGAAATATGAAAAACA 1284
 : : : : :
 Db 967 LysArgArgArgGlnGlnLysArgGlnArgGlnTyrArgLysAspLysLysLeuGlnGlnLys 986
 : : : : :
 Oy 1285 AAGATAATATATTTT---AAACCAAAATGATAAAGTTTGATGATGAGCATTTAAAAAA 1341
 : : : : :

Db 1371 HisIleSerThrLeu---ThrIleGlnLeuSerAspSerIleuIleuGlnIleuPhe 1389
1018 -----CAA 1020
Db 1390 ThrIalThrValIGluThrMetGluGluGlyIleuIleuGlnArgGluIleGluSer 1409
QY 1021 ATAAAGTAAAGTAAAGATGAATAATAGCTGTAATATGAC-----GATTCA 1065
Db 1410 LeuThrGlnGlnPheGluGluIleuAlaIleSerIleuIleuGluIleuThrIleuAsn 1429
QY 1066 TTTATAGTGAAGAAGAAATGAT-----GAAAGCTTAAAGCGAATTTAAGCCATTGTG 1119
Db 1430 ArgLeuGlnGlnGluIleuAspLeuValValAspLeuAspGlnArgGlnLeuVal 1449
QY 1120 -----CAATATGACATTTCCAAAGATGAAGAAAC----- 1149
Db 1450 SerIleuLeuGluIleuIleuGlnIleuIleuPheAspGlnMetLeuAlaGluIleuAsnIle 1469
QY 1149 ----- 1149
Db 1470 SerSerIleuIleuValAspGluArgAspArgAlaGluAlaGluAlaArgGluIleuThr 1489
QY 1150 -----ATAGCAATTTATAAAGAACTAGAAAGATTTGATAGGAAAGAAATTTAGAT 1203
Db 1490 LysAlaIleuSerLeuAlaArgAlaIleuGlnGluAlaIleuGlnAlaIleuGlnIleuGln 1509
QY 1204 GATTAGATGAAGAAATGAAGAAATCATCAGAGAAATATATCTGAAAGAAAGAAATAAAAA 1263
Db 1510 ArgThrIleuIleuMetLeuIleuAlaGluMetGluAspLeuValSerSerIleuAspAspVal 1529
QY 1264 GGAAAG-----AAATATGAAAGAAAGAAAGAT-----AAATAT 1296
Db 1530 GlyIleuAsnValIleuGlnIleuGlnIleuSerIleuArgThrIleuGlnGlnIleuGln 1549
QY 1297 TTTAAACCAATGATAAAAAGTTTGATGATAGCATATTTAAAAATATAAAAATGATAG 1356
Db 1550 MetIleuThrGlnIleuGlnIleuGlnIleuAspGlu---LeuGlnAlaIleuAspAlaIle 1568
QY 1357 -----CAGGTTAAATAGGAAAGAAAGAAATCATTAATATCATTTGTCATATATT 1407
Db 1569 LeuArgLeuGluValAsnMetGlnAla-----MetIleuSerGlnPhe----- 1582
QY 1408 GACGAGACAAATGAATTTTACAGATCGTGATGATGATATCTGAAATATATCAATAT 1467
Db 1583 -----GluArgAspLeuGlnIleuAlaArgAspGluGlnIleuGlnIleuValArgArgGln 1599
QY 1468 TTTATGAAA 1476
Db 1600 LeuLeuIleuIleu 1602
RESULT 9
S18199 myosin heavy chain - chicken (fragment)
C1Species: Gallus gallus (chicken)
C1Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 02-Feb-2001
C1Accession: S18199
R1Stewart, A.F.R.; Camoretti-Mercado, B.; Perlman, D.; Gupta, M.; Jakovic, S.; Zak, R.
J. Mol. Evol. 33, 357-366, 1991
A1Title: Structural and phylogenetic analysis of the chicken ventricular myosin heavy chain
A1Reference number: S18199; MUID:92130260; PMID:1774788
A1Accession: S18199
A1Status: Preliminary
A1Molecule type: mRNA
A1Residues: 1-1039 (578)
A1Cross-References: UNIPARC:UPI0000171256; EMBL:X59552; NID:952995; PIDN:CAA42130.1; PIR
A1Note: In the authors' translation 45-Lys is shown after residue 40, and, consequently
C1Superfamily: myosin heavy chain; myosin motor domain homology
C1Keywords: ATP, nucleotide binding, P-loop
F1402-409/Region: nucleotide-binding motif A (P-loop)
Alignment Scores: 3.37e-10 Length: 1039
Preliminary No.:

Score:	323.50	Matches:	114
Percent Similarity:	40.99%	Conservative:	143
Best Local Similarity:	22.81%	Mismatches:	187
Query Match:	13.01%	Indels:	183
DB:	2	Gaps:	27
US-09-837-344-41 (1-1482) x SI8199 (1-1039)			
QY	1 CAAGAACAACAAAAGCGATCTTGAACAAGAGACGCTGCTAAAGAA-----		45
DB	32 GTUGlmetAenSerGltueuThSerlybSArGlyLeuGluApsrGluPheAlaleu		51
QY	46 AAGTTGCAGAAACAACAAGCATTTTGA-----CAAGATAGACTTCTCTAAGAAAAG		99
DB	52 GTUleuLySlyeAspIleAspAsrleuGluIleThLeuAlaLyValJclulysGluLy		71
QY	100 -----TTACAAGAGCAAGCAAGCATTTTGAACA-----		129
DB	72 GTAlaIatThrGluAsnLyValIysAsnleuThrGluGluMetAlaThreueAspGluAen		91
QY	130 ---GAGAGACTTGCTTAAAGAAA-----TTGCAAGAA-----CAACAA-----AGC		168
DB	92 IleserLySleuThrlyeGltulysLySerSerleuGlnGluAlaHieGlnGlnValIleuAsp		111
QY	169 GATCTTGAACAACAAGAGA-----CGTGCTTAAAGAAAAGTTGCA		207
DB	112 AsrleuGlnAlaGltGluApsrLyValAsnArGlySerLeLyAlaLyValIysleuGlu		131
QY	208 GAACAACAAGCGAT-----TTAGAACAAGAGAGA-----		237
DB	132 GlnGlnValAspAsrleuGlnGlySerleuGlnGlnGlnIulysLyValArGlyMetAsrleu		151
QY	238 ---CGTGCTTAAAGAAAAGTTG-----CAAGACAACAAGCGAT		273
DB	152 GTuArGAlaLyArGlyLeuGlnGluApsrleuLyLeuThrGlnGlnuSerValIleAsp		171
QY	274 TTAGAACAAGATAGACTT---GCTAAAGAAAAGTTTACAAGACGCAAGACGCTTTTGA		330
DB	172 LeuGlnAsnAspLySleuGlnMetGlnGlnIulysLySlyeGlnPheGlnMetSer		191
QY	331 CAA-----GAGAGACGCTGCT		345
DB	192 GlnLeuAsnSerLySleIleGluApsrGlnGlnAlaIleValIleClnleuGlnIulysIle		211
QY	346 AAAGAA-----AAGTTGCAGAAACAACAAGCGATTTGAACAAGAGACGCTGCT		396
DB	212 LySgltleuGlnAlaArGltleuGlnleuGlnGlnIulysleuGlnAlaIa		231
QY	397 AAAGAAAAGTTGCAGAAACAACAAGCGATTTTGAACAAGAGACGCTTTAAAGAAAAG		456
DB	233 ArGAlaLyValGlnIulysGlnArGlySerAsrleuAlaArGlu-----LeuGlnGlu		248
QY	457 TTGCAAGAACAA-----CAAGCGATTTTGAACAAGAGAGA		492
DB	249 LeuSerGlnArGlyleuGlnIulysAlaGlyAlaIatThrAlaIaGlnleuGlnMetAsnLyS		266
QY	493 CGTGCTTAAAGAAAAGTTGCAGAAACAACAAGCGATTTTGAACA-----		537
DB	269 LybArGAlaIaGltPheleuLySleuAlaArGAsrleuGlnIulalathrlyeunHieTy		288
QY	538 -----GAGAGACGCTGCTTAAAGAAAAGTTGCAAGACAACAAGC		576
DB	289 GlnAlaIatThrAlaIaIeLeuArGlyLySbHieValAspAsrArGlyGlyMetGlyGlu		308
QY	577 GATTTGAACAACAAGAGACGCTGCTTAAAGAAAAGTTGCAAGACGCAAGAGAGATTTGA		636
DB	309 GlnLeuAspAsnleuGlnArGlyValIySglnLySleuGlnIulysGlySerGlyLeuLyS		328
QY	637 CAAAGCAAGCGCTGATTCGAAAAAAAATTTTGAAGAAAAGAAAAGCAATGGAGATATATTA		696
DB	329 MetGlnValAspAsrleuThrAlaAsnMetGlnGlnThrValLyS---GlyLybAlaAsn		347
QY	697 GCAGAGAGATTTATATGCTGCTTTAGA-----		723

```

Db      ||||| ||||| |||||
348 AlaGluValLeuThrGlyThrTyrGluIleAspHisValLeuGlnIleThrTyrLeuLeuAsp 367
Qy      724 -----ATACCACTGATAGCAATTCCTGACAGAAATGAAAGCT 759
Db      368 GluMetThrArgLeuMetAspLeuThrThrGlnIleThrTyrLeuGlnIleSerGluAsn 387
Qy      760 GGATATTATATA-----CCACATCAATCTCTTAACTCAGACACAGACAGG 807
Db      388 GluIlePheValArgGlnIleGluIleGlySerIleLeuSerGlnIleSerArgGly 407
Qy      808 AATGTAGAGATCCAGGAATA----- 831
Db      408 LysThrSerPheThrGlnGlnIleGluIleValArgGlnIleGluIleThrTyrLeu 427
Qy      832 -----TCTATATATGAA 843
Db      428 SerLysAspAlaLeuAlaIleThrAlaLeuGlnIaAlaArgHisValArgCysAspLeuLeuArg 447
Qy      844 AAAACCAATAGACAAATCTATTATCAACAAATGTTAGAGACAGAGGATATACATTAAGCA 903
Db      448 GluGlnIleTyrGluIleGluIleGlnIleValAlaLysValAspLeuGlnIleValSerLysGly 467
Qy      904 CATCTGTGA-----GAAAAGAAAGATGGTCAATAAACCAAGAA 942
Db      468 AsnAlaGluValAlaGlnIleTyrArgThrTyrGlnIleThrAspAlaIleGlnArgThrGln 487
Qy      943 CAAAAGAGATAAATCTGCTGACATA-----CAAAATCATACATTAGACAGACATA 993
Db      488 GluLeuGlnIleAspAlaLysLeuLysLeuLeuAlaArgLeuGlnIleGlnIleGlnIleAla 507
Qy      994 AATATTCTGATGTTAATGATTTTCAATAAGTAAGTATGACGATGAAATAGTGCCTGAA 1053
Db      508 GluAlaAlaAspAlaLysCysSerSerLeuGlnIleValAlaValHisValArgLeuGlnIleAsnGln 527
Qy      1054 TATGACGATTCATATAGATGATGAAAGATGATGAAAGCTTAAGCAATTTAAGCCT 1113
Db      528 GlnGluAspMetMetIleAspLeuGlnIleValAlaAsnSerAlaAlaAlaSerLeu----- 545
Qy      1114 ATTGTGCATATGACAAATTTCCAGATGAGAAACATAGCAATTTATTAAGAACTAGAA 1173
Db      546 -----AspLysLysGlnIleArgGlnPheAspLysIleIleAsn 557
Qy      1174 GATTGTAGACGAAATATGAAATTTAGATGATTTAGTAAAGCAATTAAGAAATCATCA 1233
Db      558 AspTyrLysGlnIleTyrGluIleGlnIleSerGlnIleGlnIleValAla 577
Qy      1234 GAAGAAATTTATCTGAAGAAATTAATAAAAGAAAGAAATTTGAAGAAACAAAGATTAAT 1293
Db      578 ArgSerLeuSerThrGlnIlePheLeuLeuLysAsnAlaTyrGlnIleThrLeuAspHis 597
Qy      1294 -----AATTTTAAACCAATGATTAAGTTTGTATGATGAG----- 1329
Db      598 LeuGlnIleThrLeuLysArgGlnIleAsnLysAsnLeuGlnIleGlnIleSerAspLeuThrAsn 617
Qy      1330 -----CATATTAATAAATTAATAAATGATTAAGCAG 1359
Db      618 GlnIleSerGlnIleGlnIleAsnLysAsnLeuIleGlnIleGlnIleValLys-----LysGln 635
Qy      1360 GTTAATAAGCAAAAGCAAA 1380
Db      636 ValGluGlnIleLysSerGln 642

```

RESULT 10
 A33430
 h-caldesmon - chicken
 N:Alternate names: caldesmon, smooch muscle; calmodulin- and actin-binding protein
 C:Species: Gallus gallus (chicken)
 C:Date: 27-Feb-1990 #sequence revision 27-Jun-1994 #text change 09-Jul-2004
 C:Accession: A33430; A32642; A32445; A4064; A60461; PC2003; PX0022
 R:Hayashi, K.; Kanda, K.; Kimizuka, F.; Kato, I.; Sobue, K.
 Biochem. Biophys. Res. Commun. 164, 503-511, 1989
 A:Title: Primary structure and functional expression of h-caldesmon complementary DNA.

A:Reference number: A33430; MUID:90026426; PMID:2803315
 A:Accession: A33430
 A:Molecule type: mRNA
 A:Residues: 1-771 <HA>
 A:Cross-references: UNIPROT:P12957; UNIPARC:UPI0000126E45; GB:M28417; NID:G211895; PIDN
 A:Experimental source: gizzard
 A>Note: part of this sequence was confirmed by protein sequencing
 R:Bryan, J.; Imai, M.; Lee, R.; Moore, P.; Cook, R.G.; Lin, W.G.
 J. Biol. Chem. 264, 13873-13879, 1989
 A:Title: Cloning and expression of a smooth muscle caldesmon.
 A:Reference number: A32642; MUID:89340480; PMID:2760048
 A:Accession: A32642
 A:Molecule type: mRNA
 A:Residues: 1-318,334-771 <BRY>
 A:Cross-references: UNIPARC:UPI00001713AB; GB:J04968; NID:G212656; PIDN:AAA49067.1; PID
 A>Note: the authors translated the codon GAA for residue 743 as Leu
 R:Hayashi, K.; Yamada, S.; Kanda, K.; Kimizuka, F.; Kato, I.; Sobue, K.
 Biochem. Biophys. Res. Commun. 161, 38-45, 1989
 A:Title: 35kDa fragment of h-caldesmon conserves two consensus sequences of the tropomy
 A:Reference number: A32445; MUID:89273666; PMID:2730665
 A:Accession: A32445
 A:Molecule type: mRNA
 A:Residues: 466-771 <HA2>
 A:Cross-references: UNIPARC:UPI000015624D; GB:M26684; NID:G211897; PIDN:AAA48811.1; PID
 R:Ma, A.S.; Carpenter, M.; Smillie, L.B.; Wang, J.H.
 J. Biol. Chem. 266, 19971-19975, 1991
 A:Title: Phosphorylation of caldesmon by p34(cdc2) kinase. Identification of phosphoryl
 A:Reference number: A41064; MUID:92041815; PMID:1939059
 A:Accession: A41064
 A:Molecule type: protein
 A:Residues: 597-600,678-696,711-721 <MAK>
 A:Cross-references: UNIPARC:UPI0000173E63; UNIPARC:UPI0000173E64; UNIPARC:UPI0000173E65
 R:Wang, C.L.A.; Wang, L.W.C.; Lu, R.C.
 Biochem. Biophys. Res. Commun. 162, 746-752, 1989
 A:Title: Caldesmon has two calmodulin-binding domains.
 A:Reference number: A60461; MUID:89334885; PMID:2757638
 A:Accession: A60461
 A:Molecule type: protein
 A:Residues: 2-17, 'X', 19-38,466-485 <WAN>
 A:Cross-references: UNIPARC:UPI0000173E66; UNIPARC:UPI0000173E67
 R:Hatanu, M.; Hayashi, K.; Yano, H.; Takeuchi, O.; Sobue, K.
 Biochem. Biophys. Res. Commun. 197, 145-153, 1993
 A:Title: Common structural and expression properties of vertebrate caldesmon genes.
 A:Reference number: PC2003; MUID:94071934; PMID:8250939
 A:Accession: PC2003
 A:Molecule type: DNA
 A:Residues: 74-419 <HAR>
 A:Cross-references: UNIPARC:UPI0000173E68
 R:Takeagi, T.; Yazawa, M.; Ueno, T.; Suzuki, S.; Yagi, K.
 J. Biochem. 106, 778-783, 1989
 A:Title: Amino acid sequence studies on cyanogen bromide peptides of chicken caldesmon
 A:Reference number: PX0022; MUID:90130380; PMID:2613684
 A:Accession: PX0022
 A:Molecule type: protein
 A:Residues: 462-477, 'D', 479-563,674-762, 'A', 763-771 <TRK>
 A:Cross-references: UNIPARC:UPI0000173E69; UNIPARC:UPI0000173E6A
 C:Comment: This protein plays a vital role in the regulation of smooth muscle and nonmu
 C:Comment: The binding of caldesmon to F-actin is modulated by calcium and calmodulin.
 C:Comment: Two calmodulin molecules can bind to nonoverlapping domains of each caldesmo
 C:Superfamily: caldesmon
 C:Keywords: actin binding; alternative splicing; calmodulin binding; cytoskeleton; musc
 F:1-771/Product: h-caldesmon #status predicted <HMMT>
 F:1-318,334-771/Product: h-caldesmon, alternative splice form #status predicted <HMMT>
 F:266-390/Region: 13-residue repeats
 F:511-582/Region: tropomyosin binding
 F:622-636/Region: tropomyosin binding
 F:597,682,717/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status experim
 F:688,711/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status experimen

Alignment Scores: 4.8e-10 Length: 771
 Pred. No.: 321.00 Matches: 122

Percent Similarity: 47.07% Conservative: 111
 Best Local Similarity: 24.65% Mismatches: 172
 Query Match: 12.91% Indels: 90
 DB: 1 Gaps: 21

US-09-837-344-41 (1-1482) x A33430 (1-771)

```

QY      1 CAAGAACAAACGAGATCTAGAACAGAGACGCTTAAGAAAGTTGCAAGACAA 60
DB      179 LysLysGluGluLysPserGluGluLysProLysGluValProThrGluGluLys 198
QY      61 CAAGCGATTTAGAACAGATAGACTTGTAAAGAAAGTTCAAGAGCAGCAACGAT 120
DB      199 GluValAsp-----ValAlaValGluLysSerThrAspLysGluGluVal 213
QY      121 TTAGAACAAAGAGACTTGCT-----AAAGAAAAGTTGCCAAGAA 159
DB      214 ValGluThrLysThrLeuAlaValAsnAlaGluAsnAspThrAsnAlaMetLeuGluGly 233
QY      160 CAACAAAGGATCTTGAACAGAGACGCTTAAGAAAGTTGCAAGACAAAGC 219
DB      234 GluGluSerLThrAspAlaAlaAspLysGluLysGluGluAlaGluLysGluLysGlu 253
QY      220 GATTTAGAACAAAGAGACGCTGCTAAAGAAAGTTGCAAGACAAACGATTTAGAA 279
DB      254 LysLeuGluAlaGluGluLysGluLysGluLysGluLysGluLysGluLysGlu 273
QY      280 CAAGATAGACTTGCTTAAGAAAGTTA-----CAAGAGCAGCAACGATTTA 327
DB      274 GluLysGluLysAlaGluGluGluLysGluLysAlaAlaGluGluLysGluLysAla 293
QY      328 GAACAAAGAGACGCTTAAGAAAGTTGCAAGACAAACGATTTAGAACAGAG 387
DB      294 GluGluGluLysArgAlaAlaGlu-----GluArgGluArgAlaLysAlaGluGlu 311
QY      388 AGACGTGCTTAAGAAAGTTGCAAGACAAACGATTTAGAACAGAGACTTGCT 447
DB      312 ArgLysAlaAlaGlu-----GluArgGluArgAlaLysAlaGluGluLysGluLysAla 329
QY      448 AAAAGAA-----AAGTTGCAAGACAAACGATTTAGAACAGAGACGCTGCT 498
DB      330 AlaGluGluLysArgAlaLysAlaGluGluGluLysGluLysAlaAlaGluGluLysAla 349
QY      499 AAGAA-----AAGTTGCAAGACAAACGATTTAGAACAGAGACGCTGCTA 552
DB      350 GluGluGluLysArgLysAlaAlaGluGluLysArgAlaLysAlaGluLysGluLysAla 369
QY      553 GAAGAGTTGCAAGACAAACGATTTAGAACAGAGACGCTGCT----- 600
DB      370 Glu-----GluArgGluLysArgAlaLysAlaGluGluLysGluLysAlaAlaGluGluLys 387
QY      601 -----AAGAAAGAGTTGCAAGAGCAGCAACGATTTAGAACAAAGAGGCT 648
DB      388 AlaArgLeuGluLysAlaGluLysLeuLysGluLysLys-----MetGluGluLysAla 406
QY      649 GATACGAAAAA-----AATTAGAAAGAAAGAAACATGCAATATATTAGCA 699
DB      407 GluGluGluLysAlaGluAlaAsnLeuLysArgLysGluGlu----- 420
QY      700 GAGGATTTATAGTGTGCTTAGAAATACAGATATGAACTTCATCAGAAATGAAAGCT 759
DB      421 GluAspLysGluLysAlaLysValGluLysLysGluSerLeuProGluLysLeuGlu 439
QY      760 GGATATTATATACACATCAATCTTCTTACCTCAGACACAAAGAGGAATAGTAGAGAT 819
DB      440 -----ProThrSerLysAspGluValLysAsp 449
QY      820 TCACAGAAATATCTTAATAGAAAGAAACAAATAGCAATCTTATCAACAAATGTGGA 879
DB      450 AsnLysAsp-----LysGluLysAlaProLysGluGluMetLysSerValTPAP 466
QY      880 GACGAGGAGATATACATTAAGAGACATCTTGAAGAAAGAAAGAGATGTTCAATTAACCA 939
  
```

```

DB      467 ArgLysArgGlyVal-----Pro 472
QY      940 GAACAAAGAAAGATTAATCTGCTGACATACAAATCATACATTAGACAGTAATAT 999
DB      473 GluGluLys-----AlaGluAsnGlyGluArgGluLeuThrThrProLysLeu 488
QY      1000 ---TCTGATGTTAATGATTTTCAATTAAGTATGATGAGATGAATTAAGTCTGAATAT 1056
DB      489 LysSerThrGluAsnAlaPheGlyArgSerAsnLeuLysGlyAlaAlaAsnAlaGluLys 508
QY      1057 GACGATTCATTAATAGATGAAGAAAGATGATGAA-----GACTTGACGAATTTAAG 1110
DB      509 GlySerGluLysLeuLysGluLysGluGluGluGluAlaValAlaGluLeuAspGluLys 528
QY      1111 CCTATTGCAATATGACAAAT-----TTCAAGATGAGAAACATGAGGAATTTATTA 1164
DB      529 LysArgArgGluGluLysArgLysLeuLeuGluGluGluGluLysLysGluGlu 548
QY      1165 GAACTAGAAGATTTGATAGACAAAGAAATGAAATTTAGATGATGAGAAATAGAA 1224
DB      549 GluAlaGluLysLysIleArgGluGluGluGluLysLysArgMetLysGluGluGlu 568
QY      1225 AATATCATGAGAAATTAATCTGAGAAAGAAATTAAG-----AAGAGAAAGAAATAT 1275
DB      569 ArgArgArgAlaGluAlaAlaGluLysArgGluLysValProGluAspGlyValSerGlu 588
QY      1276 GAAGAAACAAAGATTAATTAATTTAAACCAATGATTAAGTTGTATGATGAGCATATT 1335
DB      589 GluLysLysProPheLysCysPheSerProLysGluSerSerLeuLysIleGluLysArg 608
QY      1336 AAAAATATTAAGATTAAGACAGGTTTAATTAAGAAAGAAAGAA 1380
DB      609 AlaGluPhe-----LeuAsnLysSerAlaGluLys 618

RESULT 11
JC5421
smooch muscle myosin heavy chain 2 - mouse
C.Species: Mus musculus (house mouse)
C.Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C.Accession: JC5421
R.Haegawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A.Title: Molecular cloning and expression of murine smooch muscle myosin heavy chains.
A.Reference number: JC5420; MUID:97242182; PMID:9125171
A.Accession: JC5421
A.Molecule type: mRNA
A.Residues: 1-1938 <HAS>
A.Cross-references: UNIPROT:O08638; UNIPARC:UPI00002A11C; DDBJ:D85924; NID:G1945079; PI
A.Experimental source: smooch muscle
C.Comment: This protein plays a role in smooch muscle cell contraction.
C.Superfamily: myosin heavy chain; myosin motor domain homology
C.Keywords: nucleotide binding; P-loop
F.188-771/Domain: myosin motor domain homology <MMOT>
F.178-185/Region: nucleotide-binding motif A (P-loop)

Alignment Scores:
Pred. No.: 5,236-10      Length: 1938
Score: 319.00      Matches: 144
Percent Similarity: 42.02%      Conservative: 114
Best Local Similarity: 23.45%      Mismatches: 180
Query Match: 12.83%      Indels: 176
DB: 2      Gaps: 27

US-09-837-344-41 (1-1482) x JC5421 (1-1938)
QY      4 GAACAAACAAAGGATCTTGAACAGAGACGCTTAAGAAAGTTGCAAGACAA 63
DB      932 GluGluGluGluLysAspArgArgGluGluGluGluGluGluLysLysMetAlaGluGlu 951
QY      64 -----AGCGATTTAGACACAGATGACTTGCTTAAGAAAGTTACAGAG 108
DB      952 MetLeuAspLeuGluGluGluGluGluGluGluGluGluLysAlaAlaArgGluLysLeuGlu 971
  
```

QY 109 CAGCA-----AGGAT 120
: : : : :
Db 972 GluValThrAlaGluAlaLysIleValLeuGluThrAspIleLeuValMetAsp 991
QY 121 TTGAACAAGAGACTTGTCTAAAGAAAG-----TTGCAGAACCAACCAAGCATCTA 174
: : : : :
Db 992 AspGlnAsnSerIleValSerIleValArgIleValLeuGluGluValArgValSerAspLeu 1011
QY 175 -----GACCAAGAGAGACGTGCTTAA----- 195
: : : : :
Db 1012 ThrThrAsnLeuAlaGluGluGluValAlaValAsnLeuThrIleValLeuValSerIle 1031
QY 196 -----GAAAGCTTGCAGAACCAACCAAGCATTTTGAACAGAGACGTGCTTAAAGAA 249
: : : : :
Db 1032 HisGluSerMetIleSerGluLeuGluValArgLeuValSerGluValSerArgGln 1051
QY 250 AAGTTCAGAACCAACCAACGATTTTGAACAAGAT----- 285
: : : : :
Db 1052 GluLeuGluIleValLeuValSerIleValSerIleValSerIleValSerIleValSer 1071
QY 286 -----AGACTTGCTTAA-----GAAAG 303
: : : : :
Db 1072 AlaAspLeuGluAlaGluIleAlaGluLeuLeuMetGlnLeuAlaValSerGluGluGlu 1091
QY 304 TTACCAAGAGACCAAGCATTTTGAACAAG----- 336
: : : : :
Db 1092 LeuGluAlaAlaLeuAlaValSerIleValSerIleValSerIleValSerIleValSer 1111
QY 337 -----AGACGTGCTTAAAGAAAGTTGCAAGAACCAACGATTTTGAACAAGAGAGA 390
: : : : :
Db 1112 LysIleArgGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1131
QY 391 CGTGTAAAGAAAGTTTGAACAAGCAACGATTTTGAACAAGAGAGAGATTTGCT-- 447
: : : : :
Db 1132 AlaAlaArgGluValSerIleValSerIleValSerIleValSerIleValSerIleVal 1151
QY 448 AAGAAAGAGTTTGCAGAA-----CAACAAGCATTTTGAACAAGAGAGACGTGCTTAA 501
: : : : :
Db 1152 LysThrGluLeuGluLeuAspThrLeuAspSerThrAlaThrGlnGlnGluValSerGlu 1171
QY 502 GAAAG-----TTGCAGAACCAACCAAGCATTTTGAAG 534
: : : : :
Db 1172 ArgGluGlnGluValThrValLeuValValAlaLeuAspGluGluThrArgSerHisGlu 1191
QY 535 -----CAAGAGAGACGTGCTTAA-----GAAAGTTGCAAGAACCAACGAT 579
: : : : :
Db 1192 AlaGluValAlaGlnIleMetArgGlnIleValThrGlnAlaValGluGluLeuThrGluGln 1211
QY 580 TTGAACAAGAGAGACGTGCTTAAAGAAAGTTGCAAGAGACCAACGATTTTGAACAA 639
: : : : :
Db 1212 LeuGluGlnIleValSerIleValSerIleValSerIleValSerIleValSerIleVal 1231
QY 640 AGGAAGCGTGAT-----ACGAAGAAAGAAATTTTGAAG 669
: : : : :
Db 1232 GluAsnHisAspLeuAlaGluLeuValValLeuGluGlnAlaValSerGlnIleValGlu 1251
QY 670 AGAAGAAAGAGACATGAGATATTTAGCAGAGCATTTATATGCTGCTTAAAGAAATCA 729
: : : : :
Db 1252 HisValSerIleValSerIleValSerIleValSerIleValSerIleValSerIleVal 1264
QY 730 GCTATAGAACTTCATCAGAAAGTGAACGTGATATTTATATCCACATCATCTTCTTTA 789
: : : : :
Db 1265 -----GlnSerLysCys 1268
QY 790 CCTCAGGACAAAGAGGAATAGTAGATTTCCAGAGAAATATCTATATATGAAAGAAACA 849
: : : : :
Db 1269 SerAspGluValArgAlaValArgIleLeuSerAspValHisValSerGluGln---Asn 1287
QY 850 AATAGAGAAATCTTATTA-----ACAAATGTTGAAGAGCAAGAGATATATCATATAA 900
: : : : :
Db 1288 GluValGluSerValThrGluMetLeuAsnGluAlaGluGluValSerIleValSerIle 1303
QY 901 GGACATCTTGAAGAAAGAAAGATGCTTCAATTAACCAAGAAAGAAAGATTAATCT 960

Db 1304 -----IleValLeuAlaValSerValAlaSerLeuGlySerIleLeuGlnAspThrGln 1321
: : : : :
QY 961 GCTGACATACAAATATCATTTAGAGACATTAATTTCTGATGTTAATGATTTTCAAA 1020
: : : : :
Db 1322 GluLeuLeuGlnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1337
QY 1021 ATTAAGTATGATGAGATGAAATTAAGTGCGAATATGACATTCATTAATGATGAAGAA 1080
: : : : :
Db 1338 LeuArgGlnLeuGluValSerGluValSerGluValSerGluValSerGluValSerGlu 1356
QY 1081 GAAGATGATGAGACATTTAGACGAATTTAAGCCTATTGTCATATGACATTTCCAAAGAT 1140
: : : : :
Db 1357 GluAlaValSerGlnAsnLeuGluValArgHis-----ValSerThrLeuAsnIleGlnLeu 1373
QY 1141 GAAGAAACATGAGAAATTTATTAAGAACTTAAGAAATTTGATGAGAAAAATGAATTTTA 1200
: : : : :
Db 1374 SerAspSer-----LysValSerLeuGlnAspPheAlaSerThrIleGluValMet 1390
QY 1201 GATGATTTAGATGAGAAATGAGAAATGACATCGAAGAAATTAATCT-----GAAAGAA 1251
: : : : :
Db 1391 GluGluGluIleValSerValSerGluGluGluGluGluGluGluGluGluGluGluGlu 1410
QY 1252 AATATTAAGAAAGAAAGAAATATGAGAAAGAAAGAT----- 1290
: : : : :
Db 1411 LysAlaAlaAlaValThrAspValSerGluGluGluGluGluGluGluGluGluGluGlu 1430
QY 1291 -----AATTAATTTTAAACCAATGAT 1311
: : : : :
Db 1431 AspLeuValValAspLeuAspAsnGlnArgGlnLeuValSerAsnLeuGluIleValSer 1450
QY 1312 AATAGTTTATGATGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1350
: : : : :
Db 1451 LysValSer---PheAspGlnLeuLeuAlaGluGluValSerIleSerSerIleValAlaAsp 1469
QY 1351 -----GATTAAGCAGTTAATTAAGAAAGAAAGAAATTTCAATTAATCATTTGTTCAATATA 1404
: : : : :
Db 1470 GluArgAspArgAlaGluAlaGluAlaArgGluIleValSerGluValSerGluValSer 1488
QY 1405 TTGACGAGACAAATGAAATTTTACAGATCGTGAATGATTA 1446
: : : : :
Db 1489 AlaArgAlaLeuGluGluAlaLeuGluAlaValSerGluGluGluGluGluGluGlu 1502

RESULT 12
JC5420
smooth muscle myosin heavy chain 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C:Accession: JC5420
R:Hasagawa, K.; Atakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A:Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.
A:Reference number: JC5420; NUID:97242182; PMID:9125171
A:Accession: JC5420
A:Molecule type: mRNA
A:Reichers: 1-1972 (HAS-
A:Cross-references: UNIPROT:O08638; UNIPARC:UPI00000279BF; DDBJ:D85923; NID:91945077; P
A:Experimental source: smooth muscle
C:Comment: This protein plays a role in smooth muscle cell contraction.
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: nucleotide binding; P-loop
F:88-771/Domain: myosin motor domain homology <MOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)

Alignment Scores:
Pred. No.: 5,21e-10 Length: 1972
Score: 319.00 Matches: 144
Percent Similarity: 42.02% Conservative: 114
Best Local Similarity: 23.45% Mismatches: 160
Query Match: 12.83% Indels: 176
DB: 2 Gaps: 27

US-09-837-344-41 (1-1482) x JC5420 (1-1972)

[illegible]

QY 790 CCTCAGGACACGAGGGAATAGTAGAGTTCCAGAGAAATATCATATTAAGAAAAACA 849
Db 1269 SerArgGlyGluArgAlaArgAlaGluLeuSerArgPyluValHisLeuGln---Aan 1287
QY 850 AATAGAGAAATCTATTTACA-----ACAAATGTTGAGGACGAGAGGATATACATMAA 900
Db 1288 GluValGlnSerValThrGluMetLeuAenGluAlaGluGluLysAla----- 1303
QY 901 GGACATCTTGAGAAAAAGAAAGATGGTTCATTAACCAAGAACAAAGAGATAATCT 960
Db 1304 -----IleLeuValAlaLysArgValAlaSerLeuGlySerGlnLeuGlnSerThrGln 1321
QY 961 GGTGACATCAAAATCATTCATTAAGACAGATAATTTCTGATGCTTAATGATTTTCAA 1020
Db 1322 GluLeuGlnGlnGluGluThrArgGlnLysLeuAenValSerThr-----Lys 1337
QY 1021 ATAAAGTAAGTATGAGATGAGATGAATTAAGTGTCAATATGACGTTCTTAATAGTAAAGAA 1080
Db 1338 LeuArgGlnLeuGlnLysArgGluArgAenSer---LeuGlnArgGlnLeuAenGluMet 1356
QY 1081 GAAGATGATGAAGACTTAAAGCAGATTTAAAGCCTATTGTGCAATATGACAAATTTCCAAAGAT 1140
Db 1357 GluAlaLysGlnAenLeuGlnArgHis-----ValSerThrLeuAenIleGlnLeu 1373
QY 1141 GAAGAAACACATAGGAATTTTAAAGAACTAGAAAGATTTGATAGAGAAAAATGAAATTTA 1200
Db 1374 SerArgPylu-----LysLysLysLeuGlnAenPheAlaSerThrIleGluValMet 1390
QY 1201 GATGATTTTGAATGAGAAAGATAGAAATAATCATCAAGAAATATATCT-----GAAGAA 1251
Db 1391 GluGluGlnLysLysArgLeuGlnLysGlnMetGlnGluLysLeuSerGlnThrArgGlnGlu 1410
QY 1252 AAAATTAAGAAAGAAAGAAATATGAAAGAAACAAAGAT----- 1290
Db 1411 LysAlaAlaAlaLysArgPyluSerLysLeuGlnLysThrLysAenArgLeuGlnGlnLysAenPylu 1430
QY 1291 -----AATTAATTTTAAACCAATGAT 1311
Db 1431 AspLeuValAlaAspLeuArgAenGlnArgGlnLeuValSerAenLeuGlnLysLysGln 1450
QY 1312 AAAAGTTGATGATGATGATATTTAAAGAAATATTAAGAAAT----- 1350
Db 1451 LysLys---PheArgGlnLeuLeuAlaGlnLysLysAenIleSerSerLysThrAlaAsp 1469
QY 1351 -----GATTAAGCAGGTTTAATTAAGAAAGAAAGAAATTCATTAATATCATTTGTTTCATATA 1404
Db 1470 GluArgAspArgAlaGluAlaGluAlaArgLysLysGluThrLysAlaLeu---SerLeu 1488
QY 1405 TTTCACGACGACATGAAATTTTACAGATCGTGACATGACTTA 1446
Db 1489 AlaArgAlaLeuGlnGluAlaLeuAlaLysGlnGluLeu 1502

RESULT 13
A61231
Myosin heavy chain nonmuscle form A - human
N:Alternate names: cellular myosin heavy chain; myosin type 9; MNMHC-A
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Homo sapiens (man)
C:Date: 12-May-1994 #sequence, revision 14-Jul-1994 #text_change 31-Dec-2004
C:Accession: A61231, A34876, I52562, I61692
R:Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Adela, D.; Adela, J.
Circ. Res. 69, 530-539, 1991
A:Title: Human nonmuscle myosin heavy chains are encoded by two genes located on different
A:Accession: A61231, MIMD:91316803; PMID:1860190
A:Accession: A61231
A:Molecule type: mRNA
A:Residues: 1-715 51M-
A:Cross-references: UNIPROT:P35579; UNIPROT:Q9UMJ0; UNIPARC:UP1000016AD3; GB:M69180; N
R:Saev, C.G.; Myers, J.C.; Shows, T.B.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 87, 1164-1168, 1990
A:Title: Human nonmuscle myosin heavy chain mRNA: generation of diversity through alteri
A:Accession: A34876; MIMD:9018958; PMID:1967836


```
Db      1380 CysLeuGluThrAlaGluGluValLysArgLysLeuGlnLysAspLeuGlnGlyLeuSer 1399
Oy      1246 -----GAGAGAAAAATTAATAAAGAGAAAGAAATATGAAAAACAAAGCATTAATAT 1296
Db      1400 GlnArgHISGluGlnLysValAlaAlaTyrAspLysLeuGlnLysThrLysThrArgLeu 1419
Oy      1297 TTTAAACCAAAATGATTAATAAGTTGTATGAT--GAGCATATTAAATAATATATAAATGAT 1353
Db      1420 GlnGlnGlnLysLeuAspLeuValLysLeuAspHis-----GlnArgGlnSerAla 1437
Oy      1354 AAGCAGGTTAATAGCAAAAAAGAAAAATTC 1383
Db      1438 CysAenLeuGlnLysLysGlnLysLysPhe 1447

RESULT 14
H86341
hypothetical protein F9H6.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H86341
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Author: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Author: Salzman, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yi, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A66141; MUID:21016719; PMID:11130712
A:Accession: H86341
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1166 <STD>
A:Cross-references: UNIPROT:Q9SYP6; UNIPARC:UPI00000A7C33; GB:AE005172; NID:G4836896; P
C:Genetics:
A:Map position: 1

Alignment Scores:
Pred. No.: 6,82e-10 Length: 1166
Score: 317.50 Matches: 133
Percent Similarity: 41.21% Conservative: 106
Best Local Similarity: 22.93% Mismatches: 172
Query Match: 12.77% Indels: 169
Db: 2 Gaps: 24

US-09-837-344-41 (1-1482) x H86341 (1-1166)
Oy      4 GAACAACAAGCGATTTAGAACAGAGACGTCCTAAAGAAAGTTGCAAGAACACAA 63
Db      17 GlnGlnGlnLysLysLysLysSerArgSerArgSerArgSerArgSerArgLys 36
Oy      64 AGCGAT-----TTAGAACAAGAGACTGCTCTTAA----- 93
Db      37 LysAspLysGlySerGlnLysArgArgGlnLysAspArgArgLysLysValLysSer 56
Oy      94 -----GAAAAGTTACAAGAGACGAAAGCGAT 120
Db      57 SerAspSerGlnAspAspTyrAspAspArgAspAspArgGlnLysArgLysArgLys 76
Oy      121 TTAGAACAAGAGACTTCTTAAAGAA----- 147
Db      77 LysGlnArgGlnArgArgArgArgLysAspArgValLysArgArgSerGlnArg 96
Oy      148 AAGTTGCAAGCAACAAGCGATCTAGAACAGAG-----AGACGTGCT 192
Db      97 LysSerSerAspSerSerGlnAspAspValGlnGlnGlnArgGlnLysArgArgVal 116
Oy      193 AAGAAAGAAAGTTGCAAGCAACAAGCGATTTAGAACAGAGAGA----- 237
Db      117 AenGlnLys-----GlnArgGlnLysLysArgGlnLysLysArgArgArgLysAsp 133
```

```
Oy      238 CGTCTTAAGAAAGTTGCAAGAAACAAGCAAGATTTAGAACAGATAGACTTGTCTTAA 297
Db      134 ArgLysArgAspArgGlnLysArgGlnLysArgLysAspLysGlnLysArgGlnLysAsp 153
Oy      298 GAAAAGTTTACAAGACGACCAAGCGATTTAGAACAGAGAGACGTCCTAAAGAAAGTTG 357
Db      154 ArgLysArgAspGlnLysArgGlnLysArgGlnLysArgGlnLysValLysGlnLysGln 173
Oy      358 CAAGAACAAACAAGCGATTTAGAACAGAGACGTCCTAAAGAAAGTTGCAAGAACAA 417
Db      174 ArgArgGlnLysArgGlnLysArgGlnLysArgArgArgArgGlnLysGlnLysGlnLys 193
Oy      418 CAAGCGATTTAGAACAGAGACGTCCTAAAGAAAGTTGCAAGAACAAACAAGCGAT 477
Db      194 ArgArgAspArgGlnLysArgGlnLysArgSerArgGlnLysArgGlnLysSerAspAsp 213
Oy      478 TTAGAACAAGAG-----AGACGTCTTAAAGAAAGTTGCAAGAACAAACAAGCATTTA 531
Db      214 ValLysArgAspLysLysArgArgArgArgGlnLysGlnLysGlnLysGlnLysGlnLys 233
Oy      532 GAACAA-----GAGAGACGTCCTAAAGAAAGTTGCAAGAA 567
Db      234 GlnLysSerValGlnLysSerSerArgHisGlnLysSerProLysArgLysSerValGln 253
Oy      568 CAACAAGCGATTTAGAACAGAGACGTCCTAAAGAAAGTTGCAAGACGCAAGAA 627
Db      254 AspAsnGlnGlnLysLysGlnLysLysLysThrArgGlnGlnGlnLysLysLysLys 273
Oy      628 GATTTAGAACAA-----AGAAAGCGTGATACGAAAAAATTTAGAA 669
Db      274 LysLysAspGlnGlnLysLysLysLysArgArgArgArgGlnLysGlnLysGlnLys 293
Oy      670 AGAAAAAGAAACATCGAGATATATTACAGAGATTTATATGTCGTTAGAAATACCA 729
Db      294 ArgLysLysGln----- 297
Oy      730 GCTATAGAACTTCATCAGAAATGAAAGTGATATTATATACATCAATCTTCTTTA 789
Db      298 -----GlnLysLysSerLys----- 304
Oy      790 CCTCAGACCAACAGAGGAATAGTAGAGATTCCAAAGAAATTTCTATATAGAAAAACA 849
Db      305 ---GlyAspAlaAspGlnLysGlnLysProLysLysLys----- 316
Oy      850 AATAGAGATCTATTACAAACAAGCTTGAGAGCAAGAGATATCATTAAGAGCATCTT 909
Db      317 -----AlaTrpThrLeuGlnLysLysLysSerAspAspGlnLysLysPro 331
Oy      910 GAAGAAAGAA-----GATGTTCAATTAAACAGAACAAAGAAAGAA 951
Db      332 GlnGlnLysSerGlnThrGlnMetAspValAspGlnGlnLysLysProGlnLysAspGln 351
Oy      952 GAT--AAATCGTGCATACCAAAATCATCATTTAGAGACAGATTAATTTCTGATGTT 1008
Db      352 AspAlaLysMetValAspLysGlnLysGlnLysLysLysLysLysLysLysLysLys 370
Oy      1009 AATGATTTTCAATTAAGTAAGTAGAGATGAATTAAGTGCCTGAATATAGACATTCATTA 1068
Db      371 -----GlyAspGlnLysAla 374
Oy      1069 ATAGATGAAGAAAGATGATGAAAGCTTAAAGCAATTT----- 1107
Db      375 ValAspGlnGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 393
Oy      1108 -----AAGCTTATGTCATATATAGACAAATTTCCAAAGT 1140
Db      394 GlnValGlnLysPheCysAsnGlnLysLysProProLysLysLysLysLysLysLysLys 413
Oy      1141 GAAGAAACATAGCAATTTATTAAGAACTAGAAAGTTTGATGAGAAA-----AATGAA 1194
Db      414 SerLysMetAsnGlnLys-----LysGlnLysSerGlnLysAspArgProLysLysLys 431
```

QY 1195 AATTAGATGATTTAGATGAAGCATAGAAAATCATACAGAAATATTCTGAAGAAA 1254
DB 432 AAlaLeuAlaValArgIleIleGly---GluAspSerAspSerAspIyrSeGlu----- 448
QY 1255 ATAAAAAAGAAAGAAATATGAAGAAAACAAAGATATATATTTAAACCAATGATATA 1314
DB 449 -----ProlyAsnAspAspAspProSerLeuAspGluAspAsp 461
QY 1315 AGTTTGATGATGATGATGATATTAATAATATTAATAATGATAG-----CAG 1359
DB 462 GUGUluPhMetLysArgValIleYelYelThrIleValIleSerLeuValAspHis 481
QY 1360 GTTAATAGGAAAGAAAGAAATTTATTAATCATATTTCTATATTTGACGAGACAT 1419
DB 482 SerIleIleGluIyrGluProPheArgLysAsn---PheTyrIle----- 495
QY 1420 GAAATTTTACAGATCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
DB 496 GluValIleAspIleSerArgMetThrGluGluGluValAsnThrTyrArgLysGluLeu 515

RESULT 15
138055
myosin heavy chain, perinatal skeletal muscle - human
N/Content: myosin ATPase (EC 3.6.4.1)
C/Species: Homo sapiens (man)
C/Date: 17-May-1996 #sequence revision 17-May-1996 #text change 09-Jul-2004
C/Accession: J18055; JH0154; S12459; S09332; A30220; S49478
R/Jullian, E.H.; Kelly, A.M.; Pomplidou, A.J.; Hoffman, R.; Schiaffino, S.; Stedman, H.H.
Eur. J. Biochem. 230, 1001-1006, 1995
A/Title: Characterization of a human perinatal myosin heavy-chain transcript.
A/Reference number: J18055; MUID:95324556; PMID:7601129
A/Accession: J18055
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1937 <RES>
A/Cross-references: UNIPROT:P1535; UNIPARC:UPI000016ADA9; EMBL:Z38133; NID:9558668; PIR
Gene 89, 289-294, 1990
R/Karsch-Mizrachi, I.; Feghali, R.; Showe, T.B.; Letinwand, L.A.
A/Title: Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA.
A/Reference number: JH0154; MUID:90323631; PMID:2373371
A/Accession: JH0154
A/Molecule type: mRNA
A/Residues: 1-14, 'A', 16-859 <KAR>
A/Cross-references: UNIPARC:UPI000017760F; GB:Y00821
A/Experimental source: skeletal muscle
R/Bober, E.
submitted to the EMBL Data Library, January 1989
A/Reference number: S12458
A/Accession: S12459
A/Molecule type: mRNA
A/Residues: 502-1071, 'N', 1073-1250, 'DGG', 1253-1376, 'NT', 1379-1913, 'D', 1915-1937 <BOB>
A/Cross-references: UNIPARC:UPI000016A5F8; EMBL:X51592; NID:929465; PIDN:CAA35941.1; PIR
A/Experimental source: clone gEMC-F
R/Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goede, H.W.; Arnold, H.H.
Eur. J. Biochem. 189, 55-65, 1990
A/Title: Identification of three developmentally controlled isoforms of human myosin hea
A/Reference number: S09331; MUID:90235862; PMID:1691980
A/Accession: S09332
A/Molecule type: mRNA
A/Residues: 502-547, 'X', 549-617, 'X', 619-687, 'X', 689-757, 'X', 759-827, 'X', 829-897, 'X', 899-
1376, 'NT', 1379-1386, 'X', 1388-1456, 'X', 1458-1526, 'X', 1528-1596, 'X', 1598-1666, 'X', 1668-17
A/Cross-references: UNIPARC:UPI0000177610; EMBL:X51592
R/Feghali, R.; Letinwand, L.A.
J. Cell Biol. 108, 1791-1797, 1989
A/Title: Molecular genetic characterization of a developmentally regulated human perinat
A/Reference number: A30220; MUID:89234168; PMID:2715179
A/Accession: A30220
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 860-969, 'Q', 971-1246, 'H', 1248-1260, 'G', 1262-1296, 'Q', 1298-1503, 'AH', 1506-184
A/Cross-references: UNIPARC:UPI0000177611; GB:Y00821; NID:934863; PIDN:CAA68757.1; PID:G
C/Genetics:
A/Gene: GDB:MYH8

A/Cross-references: GDB:125267; OMIM:160741
A/Map position: 17pter-17p12
C/Superfamily: myosin heavy chain; myosin motor domain homology
C/Keywords: actin binding; ATP; coiled coil; hydrophobic; muscle contraction; nucleotide
F/91-769/Domain: myosin motor domain homology <MOT>
F/181-188/Region: nucleotide-binding motif A (P-loop)
F/551-588/Region: actin binding #status predicted
F/658-680/Region: actin binding #status predicted
F/842-1282/Region: S2 #status predicted
F/698,708/Active site: Cys #status predicted

Alignment Scores:
Pred. No.: 6 26e-10 Length: 1937
Score: 317.50 Matches: 156
Percent Similarity: 42.08% Conservative: 131
Best Local Similarity: 22.87% Mismatches: 184
Query Match: 12.77% Indels: 211
DB: 2 Gaps: 35

US-09-837-344-41 (1-1482) x J18055 (1-1937)

QY 4 GAAACAAGACGATCTTGAACAGAGACGCTTGAAGAA---AACTTGAAGACAA 60
DB 930 GUGUluGluGluIleAlaValIleAlaGluThrAlaLysIleValGluAspGlu 949
QY 61 CAAGCGATTTAGAACAGAT-----AACCTTCTTAA 93
DB 950 CysSerIleuLysIleAspIleAspAspLeuIleuThrIleuAlaLysValGluLys 969
QY 94 GAAAG-----TTACAAGACGACAAACGATTTGAAG 126
DB 970 GluIleHisAlaThrGluAsnLysValIleAsnLeuThrGluMetAlaGluLeuAsp 989
QY 127 CAA-----TGCAAGAA-----CAACAAGC 168
DB 990 GluThrIleAlaLysLeuSerLysGluLysLysAlaLeuGluGluThrHisGluGluThr 1009
QY 169 -----GATTCAGAACAGAGAA-----CGTCTTAAAGAAAG 201
DB 1010 LeuAspAspLeuGluIleAlaGluIleAspLysValAsnIleuThrLysAlaLysThrLys 1029
QY 202 TTGCAAGACAAACAGCGAT-----TTACAAGACAGAGA----- 237
DB 1030 LeuGluGluGluIleAlaAspLeuGluGluSerLeuGluGluGluLysLysLeuArgMet 1049
QY 238 -----CGTCTTAAAGAAAGTTG-----CAAGACAAACA 267
DB 1050 AspLeuGluArgAlaLysArgLysLeuGluGluLysLeuLysLeuAlaGluGluSerThr 1069
QY 268 AGCGATTTAGAACAGAT-----AGACTTCTTAAAGAAAGTTTACA 309
DB 1070 MetAspMetGluAsnAspLysGluGluIleuAspGluLysLeuGluLysGluPheGlu 1089
QY 310 -----GAGCAGCAAGCGATTTAGAACAGAGAGA 339
DB 1090 IleSerAsnLeuIleSerLysIleGluAspGluIleAlaValGluIleGluLeuGluLys 1109
QY 340 CGTCTTAAAGAA-----AACTTGAAGACAAACAGCGATTTAGAACAGAGAGA 390
DB 1110 LysIleLysGluLeuGluGluAlaArgIleGluGluLeuGluGluGluIleGluAlaGluArg 1129
QY 391 CGTCTTAAAGAAAGTTTGAAGACAAACAGCGATTTAGAACAGAG-----AGACTTCTT 447
DB 1130 AlaSerArgAlaLysValGluGluLysGluIleArgSerAspLeuSerArgGluLeuGluGluIle 1149
QY 448 AAAGAAAGTTTGAAGAA-----CAACAAGCGATTTAGAACAGAGAGCGT 495
DB 1150 SerGluArgLeuGluGluAlaGlyGlyAlaThrSerAlaGluIleGluLeuAsnLysLys 1169
QY 496 GCTTAAGAAAGTTTGAAGACAAACAGCGATTTAGAACAGAG----- 540
DB 1170 ArgGluAlaGluPheGluLysLeuArgArgAspLeuGluGluAlaThrLeuGluGluIleGlu 1189

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 7, 2005, 05:10:32 ; Search time 199.5 Seconds
(without alignments)

Title: US-09-837-344-41

Sequence: 1 CAAGACACAAAGCATCT.....ATATTTTATGAACTATAA 1482

Scoring table:	
BLOSUM62	
Xgapop 10.0 ,	Xgapext 0.5
Ygapop 10.0 ,	Ygapext 0.5
Fgapop 6.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 216643 beqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
	Word match	100%

```

Command line parameters:
-MODEL=frame+ n2e.model -DEV=ixh
-oq/cgna.1/USPTO_fastai_h/us09a8737344/runat_05122005_133406_29460/loop_query.Fasta_1.1677
-OB=blimpict -QPMT=specol -SUFPI=sup -MIMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-INITs=bits -START=1 -END=1 -MATRIX=bls06m2 -TRANS=human40.cd -LIST=45
-BLOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -NODE=LOCAL
-OUTPMT=pool NORH=exit HEADSIZE=5500 MINLEN=0 -MAXLEN=2000000000
-USER=0809837344 @CGC_1_1_294 @runat_05122005_133406_29460 -NCPU=6
-NO MAP -LARGEQUERY -NEG_SCORE=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEVT TIMEOUT=120 -WARN TIMEOUT=30 -THEADS=1 -KAPOP=10 -XGAPEXT=0.5 -FGAP=6
-XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -BLOP=6 -DELEXT=7

```

```
Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2478	99.7	493	2	Q25886_PLAFA	Q25886 plasmodium
2	2406.5	96.8	1509	2	Q25893_PLAFA	Q25893 plasmodium
3	2375.5	95.6	1596	2	Q81J44_PLAFA	Q81J44 plasmodium
4	1451	58.4	280	2	Q25844_PLAFA	Q25844 plasmodium
5	1450	58.3	280	2	Q25843_PLAFA	Q25843 plasmodium
6	1447	58.2	280	2	Q25847_PLAFA	Q25847 plasmodium
7	1447	58.2	280	2	Q25846_PLAFA	Q25846 plasmodium
8	1447	58.2	280	2	Q25852_PLAFA	Q25852 plasmodium
9	1445	58.1	280	2	Q25849_PLAFA	Q25849 plasmodium
10	1445	58.1	280	2	Q25851_PLAFA	Q25851 plasmodium
11	1443	58.0	280	2	Q25850_PLAFA	Q25850 plasmodium
12	1442	58.0	280	2	Q25846_PLAFA	Q25846 plasmodium
13	1437	57.8	280	2	Q25853_PLAFA	Q25853 plasmodium
14	1434	57.7	280	2	Q25854_PLAFA	Q25854 plasmodium
15	1433	57.6	280	2	Q25845_PLAFA	Q25845 plasmodium
16	1432	57.6	280	2	Q25900_PLAFA	Q25900 plasmodium

17	1430	57.5	280	2	Q25855_PLAFA	Q25855_plasmodium
18	1430	57.5	280	2	Q25868_PLAFA	Q25868_plasmodium
19	1428	57.4	280	2	Q27243_PLAFA	Q27243_plasmodium
20	1426	57.4	280	2	Q25889_PLAFA	Q25889_plasmodium
21	1366	54.9	264	2	Q9GTX5_PLAFA	Q9GTX5_plasmodium
22	770.5	31.0	318	2	Q25887_PLAFA	Q25887_plasmodium
23	468.5	18.8	978	2	Q7BARK_PLAFO	Q7BARK_plasmodium
24	451	18.1	1207	2	Q4UH84_THEAN	Q4UH84_chellieria a
25	432	17.4	905	2	Q54XN9_DICDI	Q54XN9_dicystostelli
26	429.5	17.3	3427	2	Q6YAT7_PLARE	Q6YAT9_plasmodium
27	428	17.2	1082	2	Q54UG6_DICDI	Q54J98_dicystostelli
28	422	17.0	205	2	Q54C75_DICDI	Q54C75_dicystostelli
29	422	16.9	762	2	Q5OR6A_ENTHI	Q5OR63_entamoeba h
30	419	16.9	1061	2	Q54MJ1_DICDI	Q54MJ1_dicystostelli
31	418	16.8	1575	2	Q519G4_ENTHI	Q519G4_entamoeba h
32	411	16.5	1738	2	Q51ED7_ENTHI	Q51ed7_entamoeba h
33	403	16.2	1119	2	Q54T78_DICDI	Q54t48_dicystostelli
34	396.5	16.0	1108	2	Q9NDI0_BABBI	Q9ndi0_babesia big
35	393.5	15.8	1563	2	Q75JPS_DICDI	Q75j35_dicystostelli
36	391.5	15.7	2359	2	Q81455_PLA7F	Q81455_plasmodium
37	391	15.7	1750	2	Q8XOH2_NEUCR	Q8xoh2_neutrospora
38	389	15.6	826	2	Q5VPF3_DROME	Q5vpf3_drosophila
39	383.5	15.4	1777	2	Q5AFB8_DICDI	Q5afb8_dicystostelli
40	381.5	15.3	1003	2	Q9LIX8_HNVH	Q9lix9_human herpes
41	381.5	15.3	1129	2	Q9GR71_HNVH	Q9gr71_human herpes
42	380.5	15.3	1407	1	TRHY_RAB1T	P37709_oryctolagus
43	380.5	15.3	1549	1	TRHY_SHEEP	P22773_ovis aries
44	377.5	15.2	1089	2	Q4O947_HNVH	Q4O947_human herpes
45	377	15.2	1036	2	Q9UDM3_HNVH	Q9udm3_human herpes

ALIGNMENTS

```

RESULT 1
025866_PIAFA PRELIMINARY; PRT; 493 AA.
ID Q25866_PIAFA PRELIMINARY;
AC Q25866;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DR 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Liver stage antigen-1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryote; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94267224; PubMed=7515922;
RA Flock D.A., Gras-Maes H., Lepere J., Brahimi K., Benmohamed L.,
RA Mellok S., Guerin-Marchand C., Londono A., Rahatimatala L.,
RA Meis J.F., Langsley G., Rounsillon C., Tartar A., Drulline P.;
RT "Plasmodium falciparum liver stage antigen-1 is well conserved and contains potent B and T cell determinants.";
RL J. Immunol. 153:190-204(1994).
[2]
RR NON_TER
RP NUCLEOTIDE SEQUENCE.
RA Flock D.A., Gras-Maes H., Lepere J., Brahimi K., Benmohamed L.,
RA Mellok S., Guerin-Marchand C., Londono A., Rahatimatala L.,
RA Meis J.F., Langsley G., Rounsillon C., Tartar A., Drulline P.;
RT "The Plasmodium falciparum liver stage antigen LSA-1 is well conserved and harbors major B- and T-cell epitopes.";
RL J. Immunol. 0:0-0(1994)
DR EMBL; Z30320; CAA82975.1; -; Genomic_DNA.
FT FT
SQ SEQUENCE 493 AA; 5906Z MW; FE252CC44F4696E CRC64;

Alignment Scores:
Pred. No.:          6,07e+108      Length:         493
Score:              2478.00        Matches:         493
Percent Similarity: 100.00%       Conservative:    0
Best Local Similarity: 100.00%     Mismatches:    0
Query Match:        99.68%         Indels:         0
DB:                  2             Gaps:           0

```


[illegible]

Db	1766	AlaGluTyrAspAspSerLeuIleAspGluGluGluAspAspGluAspLeuAspGluPhe	1765
Qy	1108	AAGCCTATTGTGCAATATGACCAATTTTCCAAAGATGAGAAAAACATAGCAATTTATTAAGAA	1167
Db	1786	LysProIleValIeGlnTyrAspAsnPheGlnAspGluGluAsnIleGlyIleTyrLysGlu	1805
Qy	1168	CTAGAGATTGTGATAGGAAAAATGAAAATTTTAGATGATTTTAGATGAAAGAAATTTGAAAAA	1227
Db	1806	LeuGluAspLeuIleGluTyrAsnGluAsnLeuAspAspLeuAspGluGlyIleGluLys	1825
Qy	1228	TCATCAGAGAAATATATCTGGAAGAAAAAATATAAAAAGAAAGAAATATGAAAAACAAG	1287
Db	1826	SerSerGluGluLeuSerGluGluGluLysIleLysGluGlyLysTyrGluLysPheLys	1845
Qy	1288	GATATATATTTTAAATCCAAATGATMAAAAGTTTGTATGATGAGCATATTTAAAAATATATAA	1347
Db	1846	AspAsnAsnProPheLysProAsnAspLysSerLeuTyrAspGluIleHisIleLysTyrLys	1865
Qy	1348	AATGATTAAGCAGTTTAATATAGGAAAAAGAAAAATTCTATAAATCATTTGTTTCATATATTT	1407
Db	1866	AsnAspLysGluValIeAsnLysGluLysGluLysPheIleLysSerLeuPheHisIlePhe	1885
Qy	1408	GACGAGAGCAATGAAATTTTTCACAGTGTGGATGATCTGTGAAGATATTAACATAATAT	1467
Db	1886	AspGluAspAsnGluIleLeuGlnIleValAspGluLeuSerGluAspIleIleHisLysTyr	1905
Qy	1468	TTTATGAAACTA 1479	
Db	1906	PheMetLysLeu 1909	
RESULT 3			
Qy	OB1J44_PLAF7		
ID	OB1J44_PLAF7 PRELIMINARY; PRT; 1596 AA.		
AC	OB1J44		
DT	01-MAR-2003 (TREMBLrel. 23, Created)		
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	Liver stage antigen, putative.		
GN	ORFNames=PF10.0356;		
OS	Plasmodium falciparum (isolate 3D7).		
OC	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=36329;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;		
RA	Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,		
RA	Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,		
RA	Eisen J.A., Rutherford K., Salzberg S.L., Craig S.,		
RA	Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,		
RA	Perera M., Allen J., Selengut J., Haft D., Mather M.W., Valdivia A.B.,		
RA	Martin D.M.A., Fairlamb A.H., Frumholz M.J., Roes D.S., Ralph S.A.,		
RA	McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,		
RA	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,		
RA	Frazer C.M., Barrcll B.G.;		
RT	"Genome sequence of the human malaria parasite Plasmodium		
RT	falciparum."		
RL	Nature 419:4836-511(2002).		
SD	EMBL, AE014836; AAN35553.1; -, Genomic DNA.		
SD	SEQUENCE 1596 AA; 191691 MW; C0B3A2CC0F5D60AC CRC64;		
Alignment Scores:			
	Pred. No.:	2,82e-103	Length: 1596
	Score:	2375.50	Matches: 478
	Percent Similarity:	94.71%	Conservative: 5
	Best Local Similarity:	93.73%	Mismatches: 10
	Query Match:	95.56%	Indels: 17
	DB:	2	Gaps: 1
Qy	US-09-837-344-41 (1-1482) x OB1J44_PLAF7 (1-1596)		
Db	1087 GlnGlyGlnGlnSerAspLeuGlnGlnGlnGluGluAlaGlyGluLysGluGlnGlnGlyGln		1106
	1 CAAGAACACAAAGGAGTCTTGGAAACAGACAGACGTGCTAAAGAAAAGTTGCAGAACAA		60

Db 41 G1YrTyrlleProH1eGlnSerSerLeuProGlnAhpAenAxlGlyAenSerArgAsp 60
QY 820 TCCAGAGAAATATTTATTAATGAAAAAACAATGAGAAATTTATTAACAACAAATGTTGAA 879
Db 61 SerlyseGlnleSerlleleGlnlyeThrsAenAxlGlnleSerlleThrsAenValGln 80
QY 880 GGAGCAAGGATATACATAAGCAATCTTGAGAAAAAGAAAGATGGTCAATTAACCA 939
Db 81 GlyArgArgAenAxlleH1eGlnlyeGlnleGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 100
QY 940 GAACAAAAGAGATTAATCTGCTGACATCAAAATCATACATTAGAGACAGTAATATTT 999
Db 101 GlnGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlye 120
QY 1000 TCTGATGTTAATGATTTTCAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1059
Db 121 SerAhpAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 140
QY 1060 GATTCATTAATAGATGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119
Db 141 AhpSerleuAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 160
QY 1120 CAATATGACAAATTTCCAAATGAGAAACAATAGAAATTTATTAAGAACTAGAAATTTG 1179
Db 161 GlnTyrlleAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 180
QY 1180 ATAGAGAAAAATGAAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239
Db 181 l1eGlnlyeAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 200
QY 1240 TTATCTGAAGAAAAATTAAGAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTT 1299
Db 201 LeuSerGlnGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 220
QY 1300 AAACCAATGATTAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
Db 221 LysProAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 240
QY 1360 GTTAATAGAGAAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTT 1419
Db 241 ValAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 260
QY 1420 GAAATTTTACGATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
Db 261 GlnleGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 280

RESULT 5

Q25843_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25843;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1; falciparum.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=6065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povo W.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates."
RL Mol. Biochem. Parasitol. 71:291-294(1995).
FR EMBL; L40884; AAB59230.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32926 MW; 467080F32FAEAD33 CRC64;
Alignment Scores: 5.15e-60 Length: 280

Score: 1450.00 Matches: 279
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 0
Query Match: 58.33% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-41 (1-1482) x Q25843_PLAFA (1-280)

QY 640 AGGAGGCTATACGAAAAAATTTAGAAAAAGAAAGCAATGAGATATTTAGCA 699
Db 1 ArgLysAlaAhpThrsLysAenLeuGlnAxlGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 20
QY 700 GAGATTTATATGATCGTTTATGAAATACAGATATGAACTTCCATGAGAAATGAACTG 759
Db 21 GlnAhpLeuTyrlleAxlleGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 40
QY 760 GATATTATATACACATCAATCTTCTTACCTGAGACAAAGAGAAATGATGATGAT 819
Db 41 G1YrTyrlleProH1eGlnSerSerLeuProGlnAhpAenAxlGlyAenSerArgAsp 60
QY 820 TCCAGAGAAATATTTATTAATGAAAAAACAATGAGAAATTTATTAACAACAAATGTTGAA 879
Db 61 SerlyseGlnleSerlleleGlnlyeThrsAenAxlGlnleSerlleThrsAenValGln 80
QY 880 GGAGCAAGGATATACATAAGCAATCTTGAGAAAAAGAAAGATGGTCAATTAACCA 939
Db 81 GlyArgArgAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 100
QY 940 GAACAAAAGAGATTAATCTGCTGACATCAAAATCATACATTAGAGACAGTAATATTT 999
Db 101 GlnGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlye 120
QY 1000 TCTGATGTTAATGATTTTCAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1059
Db 121 SerAhpAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 140
QY 1060 GATTCATTAATAGATGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119
Db 141 AhpSerleuAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 160
QY 1120 CAATATGACAAATTTCCAAATGAGAAACAATAGAAATTTATTAAGAACTAGAAATTTG 1179
Db 161 GlnTyrlleAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 180
QY 1180 ATAGAGAAAAATGAAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239
Db 181 l1eGlnlyeAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 200
QY 1240 TTATCTGAAGAAAAATTAAGAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTT 1299
Db 201 LeuSerGlnGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 220
QY 1300 AAACCAATGATTAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
Db 221 LysProAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 240
QY 1360 GTTAATAGAGAAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTT 1419
Db 241 ValAenAxlleH1eGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 260
QY 1420 GAAATTTTACGATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
Db 261 GlnleGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGlnlyeGln 280

RESULT 6
Q25847_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25847;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;

[illegible]

QY	1360	GTATATATAGAGAAAAAGAAAAATTCATATAATCATTTGTTTCATATATTTGACGGAGACAAT	1413
Db	241	ValanysgsluysgsluyspneilelyserieuhehstillepaeBgIyaspn	260
QY	1420	GAATTTTACGATCGTGATGAGTATTCGAGATATTAATAATTTTATGAACATA	1479
Db	261	GulileuGnilleValapBgIuleuSeGluaspIlethrltyrtyrPhemeUlyleu	280
RESULT 7			
Q25848_PLAFA			
ID	Q25848	PLAFA PRELIMINARY;	PRT; 280 AA.
AC	Q25848		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-OCT-2002	(TREMBLrel. 22, Last annotation update)	
DE	Liver stage-specific antigen 1 (Fragment).		
GN	Name=LSA-1;		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=5833;		
RN	(1)		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6861(95)00069-D;		
RA	Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,		
RA	Hawley W.A., Collins W.E., Lal A.A.;		
RT	"Sequence variations in the non-repetitive regions of the liver stage-		
RT	specific antigen-1 (LSA-1) of Plasmodium falciparum from field		
RT	isolates.";		
RL	Mol. Biochem. Parasitol. 71:291-294(1995).		
DR	EMBL; LA0889; AAC41599.1; -; Genomic_DNA.		
FT	NON TER		
SQ	SEQUENCE	280 AA; 32939 MW; E77080365FAM061	CRC64;
Alignment Scores:			
Pred. No.:	7.1e-60	Length:	280
Score:	1447.00	Matches:	279
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.64%	Mismatches:	0
Query Match:	58.21%	Indels:	0
DB:	2	Gaps:	0
US-09-837-344-41 (1-1482) x Q25848_PLAFA (1-280)			
QY	640	AGGAAGCGTGTATGAGAAAAAAATTTGAAAGAAAAAGAAACATGAGATATATTAGA	699
Db	1	ArgylsalaaaprrnlvlysaenleGluTgIyulysGluIslgIyaspIleleuAla	20
QY	700	GAGGATTTATATGCGTCTTTGAAATACCAAGCTATAGAACTTCATCAGAAATGAACGT	759
Db	21	GluaspIeulTgIyulTgIleGluIleProlaIleGluIleuPProserGluaspnGluArg	40
QY	760	GGATTTATATATACCAATCAATCTTTCTTTTACCTGAGACAACAAGGAAATAGTAGAT	819
Db	41	GlyTyrTyrIleProlaIleGlnSerSerleuProlaIlePaspnAdgIyaspSerArgasp	60
QY	820	TCCAAGGAATATCTATATATAGAAAAAACAATGAGAACTTATTAACAACAATGTGAA	879
Db	61	SerlysgIuIleSerIleIleGluIlylThaspnTgIuSerIleThaspnValGlu	80
QY	880	GGACGAGGATATATACATATAAGACATCTTGAGAAAAAGAAAGATGGTTCATATAACA	939
Db	81	GlylTgArgaBpIleHleHleGlyHleuGluIyulysylvaBpGlySerIleIyPPro	100
QY	940	GAACAAAAGAAGATAATCTGCTGACATACAAAATCATACATTAGAGACAGTAAATATT	999
Db	101	LysGlnlysgIuaspIlySerAlaaspIleGlnaspnsthrlleuGluThValaspnIle	120
QY	1000	TCTGATGTTATATGATTTTCAATAATAGTAGATATGAGTGAATAATAGCTGATATAGAC	1059
Db	121	SeraspValaspnBpGlnIleSerlysrtyGluaspGluIleSerAlaGluTyrasp	140
QY	1060	GATTCATTTATATAGATGAAGAAGATGATGAACACTTAGACGAATTTAAAGCTATTGTG	1119

Db	141	AspSerLeuIleAspGluGluGluAspArgIuAspLeuAspGluPheLeuProIleVal	160
Qy	1120	CAATATGACAAATTTCCAGATATGAAGAAACAATAGCAATTTATTAAGAAGCTAGAGATTGG	1179
Db	161	GlnTyrAspAsnPheGlnAspGluGluAsnIleGlyIeTyrLysGluLeuGluAspLeu	180
Qy	1180	ATAGAGAAAAATGAAAAATTTAGATGATTAGATGAAGAAATAGAAAAATTCATCAGAGAGA	1239
Db	181	IleGluLysAsnGluAsnLeuAspPheLeuAspArgIuGlyIleGluLysSerSerGluGlu	200
Qy	1240	TTATCTGAAGAAAAATATAAAAAAGAAAGAAATATGAAAAACAAGATATATATTTT	1299
Db	201	LeuSerGluGluLysIleLysLysGlyLysIleTyrGluLysThrLysAspAsnAsnPhe	220
Qy	1300	AAACCAAAATGATAAAAATTGTATGATGACATATTAATAAATATTAATAATGATTAACAG	1355
Db	221	LysProAsnAspLysSerLeuTyrAspGluHisIleLysLysTyrLysAsnAspArgGln	240
Qy	1360	GTTAATTAAGAAAAAGAAAAATTCATAAATCATATGTTTCATATATTTGACGGAGACAT	1419
Db	241	ValAsnLysGluLysGluLysPheIleLysSerLeuPheHisIlePheAspGlyAspAsn	260
Qy	1420	GAATATTTACAGATCTGCGATGACGTTATCTGAAGATATTAACATAATATTTTATGAACAATA	1479
Db	261	GluIleLeuGlnIleValAspGluLeuSerGluAspIleTyrLysTyrPheLeuTyrLeu	280

RESULT 8

Q25852	PLAFA
ID	Q25852_PLAFA PRELIMINARY; PRT: 280 AA.
AC	Q25852_
DT	01-NOV-1996 (TREMBLrel. 01, Created)
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE	Liver stage-specific antigen 1 (Fragment).
CN	Name=LSA-1;
OS	Plasmodium falciparum.
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CX	NCBJ_Taxid=5833; [1]
RN	NUCLEOTIDE SEQUENCE.
RP	MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RX	Yang C., Shi Y.P., Udhayakumar V., Alperez M.P., Povoa M.M., Hawley W.A., Collins W.E., Lal A.A.;
RA	"Sequence variations in the non-repetitive regions of the liver stage-
RT	-specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT	isolates";
RL	Mol. Biochem. Parasitol. 71:291-294 (1995).
DR	EMBL, L40834, AA859233.1; -; Genomic DNA.
FT	NON TER 1
SQ	SEQUENCE 280 AA; 32927 MW; E6648F85FBBA245 CRC64;

Alignment Scores:

Pred. No.:	7.1e-60	Length:	260
Score:	1447.00	Matches:	278
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	99.29%	Mismatches:	0
Query Match:	58.21%	Indels:	0
DB:	2	Gaps:	0

US-09-837-344-41 (1-1482) x Q25852 PLAFPA (1-280)

QY	640	AGCAAGCGCTGATACGAAAAAAAAATTTAGAAAGCAAAAAAGCAACATGGAGATATATTGCA	639
Db	1	ArglyValAlaSprhTrlySlybAaenLeGluArgLyslySglIuhIsgIyAaSprValLeuAla	20
QY	700	GAGGATTTATATGCTGCTTTAGAAATATCCAGCTATATAGAATTCCATACGAAATGAACGT	75
Db	21	GluAspLeuTrlycIyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArg	40
QY	760	GGATATATATACCAATCAATCTTCTTACCTCAGACAAACAGAGGAATATAGAGAT	819
Db	41	GlyTyrTyrIleProhIsgInSerSerLeuProGluAlaSprAsnArgIyAaenSerArgAsp	60

OY	820	TCCAAGAAATATCTTAATATGAAAAACAATTAAGAAATCTATTACACAAAGTTGAA	879
Db	61	SerTysGluLeSerIleIleGluTyrThrAsnArgGlnSerIleThrTrpAsnValGlu	80
OY	880	GGACGAAGGCGATATACATTAAGACACCTTTGAAGAAAAAGAAAGATGGTTCATTAACCA	939
Db	81	GlyArgThrArgPheIleHisGlyGlyHisIleGluGluTyrLeuValAspGlySerIleLeuPro	100
OY	940	GAACAAAGAAGATTAATCTGCTGCACATACAAATCATACATTAGACACGTAAATAT	999
Db	101	GluGluTyrGluAspLeuTyrSerTrpAlaAspIleGluAsnHisThrLeuGluTrpValAsnIle	120
OY	1000	TCGTAGTGAATGATTTTCCAAATAAGTAAGATAGAGATAAGATAAGGCGCAATTAGAC	1059
Db	121	SerSerPalaAsnAspPheGlnIleSerLeuTyrGluAspGluIleSerTrpAlaGluTyrAsp	140
OY	1060	GATTCATTAAATAGATGAAGAAGAGATGATGAAGACTTAGACGAATTTAAGCCTATTGTG	1119
Db	141	AspSerLeuIleAspGluGluGluAspArgLeuAspLeuAspGluPheLeuVal	160
OY	1120	CAATATGACATTTTCCAAAGATGAAGAAAACATAGAAATTTATTAAGAACTTAGAAGATTG	1179
Db	161	GlnTyrAspAsnPheGlnAspGluGluAsnIleGlyIleTyrLeuValGluLeuGluAspLeu	180
OY	1180	ATAGAGAAAAATGAAAAATTTTAGATGATTTAGATANAAGAAATAGAAAAATCATCAGAAGAA	1239
Db	181	IleGluTyrAsnGluIleAsnLeuAspAspLeuAspGluGlyIleGluTyrSerSerGluGlu	200
OY	1240	TTATCTGAAGAAAAATTAATAAGAAAGAAAGAAATATGAAAAACAAGGATTAATATTTT	1299
Db	201	LeuSerGluGluTyrSileTyrSylGlyLeuTyrGluTyrThrLeuAspAsnAsnPhe	220
OY	1300	AAACCAATGATGTAATAAGTTGTATGATGAGACATTTTAAAAAATTAATAAAATGATTAACAG	1359
Db	221	LysProAsnAspLysSerLeuTyrAspGluHisIleLeuSylTyrLeuAsnAspLeuGln	240
OY	1360	GTTAATAAGAAAAAGAAAAATTCATTAATCATTTGTTCTATATATTTTGACGAGACAAAT	1419
Db	241	ValAsnLeuGlyTyrSylGluTyrPheIleLeuSerLeuPheHisIlePheAspGlyAspAsn	260
OY	1420	GAATTTTACGATCGCGGATGATTTTCTGAAGATTAATTAATCTAAATTTTATGAAGACTA	1479
Db	261	GluIleLeuGlnIleValAspGluLeuSerGluAspIleThrTyrTyrPheMetLeuLeu	280

RESULT 9

ID	025849_PLAFA	PRELIMINARY;	PRT;	280 AA.
AC	025849_			
DT	01-NOV-1996	(TRENBLER). 01.	Created)	
DT	01-NOV-1996	(TRENBLER). 01.	Last sequence update)	
DT	01-OCT-2002	(TRENBLER). 22.	Last annotation update)	
DE	Liver stage-specific antigen 1 (fragment).			

OS Plasmodium f

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;

RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-replicative regions of the liver stage
RT specific antigen-1 (USA-1) of *Plasmodium falciparum* from field
RT isolates";
RT Mol. Biochem. Parasitol. 71:229-234(1995).
DR EMBL/ L40890; AAC41600.1; -; Genomic DNA.

SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEA9CF CRC64.

Alignment scores:	
Pred. No.:	8.8e-60
Score:	1445.00
	Length: 280
	Matches: 279

Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 0
Query Match: 58.13% Indels: 0
DB: 2 Gaps: 0
US-09-837-344-41 (1-1482) x Q25851_PLAFA (1-280)

QY 640 AGAAGGCTGATACGAAAAAATTTAGAAAAAGAAACATGAGATATATTACCA 699
DB 1 ArglysaLaApPThLySaSnLeuGluArgLySvGluHISGlyAspIleLeuAla 20
QY 700 GAGATTTATATGTCGTTTAAAGAAATACAGCTATAGAACTTCATTCAGAAATGAAAGT 759
DB 21 GluApleuTyrgLyArgLeuGluIleProAlaIleGluLeuProSerGluSnnIuArg 40
QY 760 GGATATTATATACCAATCAATCTTCTTACTCAGGACACAGAGGGAATGTAGAGAT 819
DB 41 GlyTyTyTyIleProHISerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
QY 820 TCCAGAAATTTCTATATAGAAAAACAATAGAGATCTTATACAAATGTTGAA 879
DB 61 SerLySgIuIleSerIleIleGluTySthraAsnArgIuSerIleThrThraSnnValGlu 80
QY 880 GGACGAGGATATACATTAAGACATCTTGAAAGAAAGAAAGATGCTCAATAAACCA 939
DB 81 GlyArgAspPleHisIleSvGlyHISLeuGluGluLySvAspGlySerIleLeuPro 100
QY 940 GAACAAAAAGAAATTAATCTGCTGACATACAAATCATACATTAAGACAGTAAATTT 999
DB 101 GluGluLySgIuAspLySerIleAspIleGlnAsnHISThrLeuGluThrValAsnIle 120
QY 1000 TCTGATGTTAATGATTTTCAAAATTAAGTATAGAGAAATAGCTGAATATGAC 1059
DB 121 SerAspValAsnAspPheGlnIleSerLyTyrgIuAspGluIleSerIleGluTyArgP 140
QY 1060 GATTCATTATAGATCAAGAAAGAAAGATGAAAGACTTGAAGCAATTTAAGCTATTGTC 1119
DB 141 AspSerLeuIleAspGluGluGluAspAspGluAspGluPheLeuAspGluPheLeuVal 160
QY 1120 CAATATGACAAATTTCCAGATGAGAAAGAAACATAGAAATTTATTAAGAAATTTG 1179
DB 161 GlnTyArgAspAsnPheGlnIleAspGluGluAsnIleGlyIleTyrgIuLeuGluAspLeu 180
QY 1180 ATAGAGAAAAAGAAATTTAGATGCTTTAGATGAAAGAAATGAAAAATCATCAGAAAGAA 1239
DB 181 IleGluTySAsnGluSnnLeuAspAspLeuAspGluGlyIleGluTySerSerGluGlu 200
QY 1240 TTATCTGAAGAAAAATTAAGAAAGAAATATGAAAAACAAGATATATATTT 1299
DB 201 LeuSerGluGluLySvIleLySvGlyTyrgIuLySvThrLySAspAsnAspPhe 220
QY 1300 AAACCAATGATTAAGATTGTAATGATGACATATTAATAATATTAATAATGATAAGCG 1359
DB 221 LysProAsnAspLySerLeuTyArgPgluTyIleLySvIleTyrgIuSnnAspLySgln 240
QY 1360 GTTAAATAGGAAAAAGAAATTTCAATTAATCTTGTTCATATATTTTGAAGAGCAAT 1419
DB 241 ValAsnLySgIuLySvPheIleLySvSerLeuPheHisIlePheAspGlyAspAsn 260
QY 1420 GAAATTTTACAGATCGTGATGAGTATCTGAAGATATACATAATATTTTGAAGAACTA 1479
DB 261 GluIleLeuGlnIleValAspGluLeuSerGluAspIleThrLySvTyrgIuTherLeuSleu 280

RESULT 10
Q25851 PLAFA PRELIMINARY: PRT: 280 AA.
AC Q25851:
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN NameLSA-1;
OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40893; AAC41603.1; -; Genomic_DNA.
FT NON TER
SQ SEQUENCE 280 AA; 32927 MW; 48706EFDCE40AD3 CRC64;

Alignment Scores:
Pred. No.: 8.8e-60 Length: 280
Score: 1445.00 Matches: 278
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.29% Mismatches: 0
Query Match: 58.13% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-41 (1-1482) x Q25851_PLAFA (1-280)

QY 640 AGAAGGCTGATACGAAAAAATTTAGAAAAAGAAACATGAGATATATTACCA 699
DB 1 ArglysaLaApPThLySaSnLeuGluArgLySvGluHISGlyAspIleLeuAla 20
QY 700 GAGATTTATATGTCGTTTAAAGAAATACAGCTATAGAACTTCATTCAGAAATGAAAGT 759
DB 21 GluApleuTyrgLyArgLeuGluIleProAlaIleGluLeuProSerGluSnnIuArg 40
QY 760 GGATATTATATACCAATCAATCTTCTTACTCAGGACACAGAGGGAATGTAGAGAT 819
DB 41 GlyTyTyTyIleProHISerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
QY 760 GATTCATTATAGATCAAGAAAGAAAGATGAAAGACTTGAAGCAATTTAAGCTATTGTC 819
DB 41 GlyTyTyTyIleProHISerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
QY 820 TCCAGAAATTTCTATATAGAAAAACAATAGAGATCTTATACAAATGTTGAA 879
DB 61 SerLySgIuIleSerIleIleGluTySthraAsnArgIuSerIleThrThraSnnValGlu 80
QY 880 GGACGAGGATATACATTAAGACATCTTGAAAGAAAGAAAGATGCTCAATAAACCA 939
DB 81 GlyArgAspPleHisIleSvGlyHISLeuGluGluLySvAspGlySerIleLeuPro 100
QY 940 GAACAAAAAGAAATTAATCTGCTGACATACAAATCATACATTAAGACAGTAAATTT 999
DB 101 GluGluLySgIuAspLySerIleAspIleGlnAsnHISThrLeuGluThrValAsnIle 120
QY 1000 TCTGATGTTAATGATTTTCAAAATTAAGTATAGAGATGAATATAGCTGAATATGAC 1059
DB 121 SerAspValAsnAspPheGlnIleSerLyTyrgIuAspGluIleSerIleGluTyArgP 140
QY 1060 GATTCATTATAGATGAGAAAGAAAGATGAGAACTTGAAGCAATTTAAGCTATTGTC 1119
DB 141 AspSerLeuIleAspGluGluGluAspAspGluAspGluPheLeuAspGluPheLeuVal 160
QY 1120 CAATATGACAAATTTCCAGATGAGAAAGAAACATAGAAATTTATTAAGAAATTTG 1179
DB 161 GlnTyArgAspAsnPheGlnIleAspGluGluSnnIleGlyIleTyrgIuLeuGluAspLeu 180
QY 1180 ATAGAGAAAAAGAAATTTAGATGCTTTAGATGAAAGAAATGAAAAATCATCAGAAAGAA 1239
DB 181 IleGluTySAspGluSnnLeuAspAspLeuAspGluGlyIleGluTySerSerGluGlu 200
QY 1240 TTATCTGAAGAAAAATTAAGAAAGAAATATGAAAAACAAGATATATATTT 1299
DB 201 LeuSerGluGluLySvIleLySvGlyTyrgIuLySvThrLySAspAsnAspPhe 220
QY 1300 AAACCAATGATTAAGATTGTAATGATGACATATTAATAATATTAATAATGATAAGCG 1359
DB 221 LysProAsnAspLySerLeuTyArgPgluTyIleLySvIleTyrgIuSnnAspLySgln 240

QY	1360	TTTAATTAAGGAAAGGAAATTCCTAAATCACTTGTTCATATATTGACGAGACAT	1413
Db	241	ValbndvysGulvysGulvyspshellelysserleupnehsillepeaspglyaspsn	260
QY	1420	GAATTTTACAGATCGTGTGATGATTAATCTCAAGATATACATAATATTTATGAACATA	1479
Db	261	GulileuGulnillevalaspgulvleusserGulvleupliethrlystyRphenellvleu	280
RESULT 11			
	Q25850_PLAFA		
ID	Q25850	PLAFA PRELIMINARY;	PRT; 280 AA.
AC	Q25850		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-OCT-2002	(TREMBLrel. 22, Last annotation update)	
DE	Liver stage-specific antigen 1 (fragment).		
GN	Name=LSA-1;		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=5633;		
RP	[1]		
RP	NCBI_EPIRIDE SEQUENCE.		
RX	MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6651(95)00069-D;		
RA	Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,		
RA	Hawley W.A., Collins W.E., Lal A.A.;		
RT	"Sequence variations in the non-repetitive regions of the liver stage-		
RT	specific antigen-1 (LSA-1) of Plasmodium falciparum from field		
RL	isolates.";		
RL	101.		
DR	Mol. Biochem. Parasitol. 71:291-294(1995).		
FT	EMBL; L40891; AAC41601.1; -, Genomic_DNA.		
SO	SEQUENCE	280 AA;	32989 MW; E36F812CFPABED8 CRC64;
Alignment Scores:			
Pred. No.:	1,096-59	Length:	280
Score:	1443.00	Matches:	279
Percent Similarity:	99.64%	Conservative:	0
Best Local Similarity:	99.64%	Mismatch:	1
Query Match:	58.05%	Indels:	0
DB:	2	Gaps:	0
US-09-837-344-41 (1-1482) x Q25850_PLAFA (1-280)			
QY	640	AGGAAGCGTAGACGAAAAAAATTTAGAAAAAGAAAAAGAACATGAGATATATTAGCA	699
Db	1	ArglybvalaapbthrlylyabamleucGulvlybysGulnlsGlyaspgilleualla	20
QY	700	GAGGATTTATATGCGCTTTAGAAATACCAGCTATAGAACTTCCATCAGAAAAAGACGT	759
Db	21	GlvbpbleuTyrglyarGlvleuGulvleProlaaileGlvleuProSeGlvabnGlvayg	40
QY	760	GGATTTATTAATACCCATCATCTCTTACCTCAGACCAACAGAGGAATATAGATAGAT	819
Db	41	GlylyTyTllepnohagInserserleuproGlvabnabnagGlylyasbserThgasp	60
QY	820	TCCAAGAAATATCTATATAGAAAAAAACAATAGAGATCTATTAACAACATGTTGAA	879
Db	61	SerlyseGulileserilleGlvlyerThraenatGlvserilleThrThraValGlv	80
QY	880	GGACGAAGGATATCATTAAGAGACATCTTGAAGAAAAAGAAAGATGCTTCAATAAACCA	939
Db	81	GlyarGatAapRillehlyeglyVhlsleuGulvlyblyaspglyserilleyspro	100
QY	940	GAAACAAAAGAGATTAATCTGCTGACATTAACAAATCTATACATTAAGACAGTAAATAT	999
Db	101	GlvnglvlybysGlvabpLyserbshlaapRilleGlvlnbnhlsrThleGlvThraValbnlle	120
QY	1000	TCTGATGTTAATGATTTTCAATATAGATATAGATATAGATGTAATATGCTGAATATGAC	1059
Db	121	SerabpValaenabpRheGlvlnlseserlybTyrglvabpGulvleaserAlGlvlyaysp	140
QY	1060	GATTCATTAATATGATGAAGAAAGATGATGAAGCTTAAGACGAATTTAAGCTTATGCTG	1119

QY	1120	CAATATGCAATTTTC	CAAGATGAGAAACCACTAGCAATTTTAAAGAACTTGAAAGATTTG	1175
DB	161	GIATYrAspaenPhaeGInAspGIuGluAsnIleGIlyIerYrLysGIuLeuGluAspLeu		180
QY	1180	ATTAGAGAAAAATGAAAAATTGATGATTTTAAATGAGAAATGAGAAAAATCATCTCAGAAACA		1235
DB	181	IlleGIuLysAsnGIuAsnLeuAspAspLeuAspGIuGIlyIleGIuLysSerSerGIuGIu		200
QY	1240	TTATCTGAGAAAAATTAATAAAAAAGAAAGAAATATGAAAAACAAAGATTAATTAATTTT		1295
DB	201	LeuSerGIuGIuLysIleYrLysGIlyYrLysSerYrGIuLysrIthIyAspTyrAsnPh		220
QY	1300	AAACCAATGATTAATAAGTTGTTATGATGACATATTTAAAAATATTAATAATGATPACAG		1355
DB	221	LysProAsnAspLysSerLeuTyrAspGIuHsIleYrLysIleYrLysAsnAspLysGIu		240
QY	1360	GTTAATTAAGAAAAAGAAAAATTCATATAAATCATTTGTTCAATATTTTGAAGGACACAT		1415
DB	241	ValAsnLysGIuLysGIuLysPheIleYsSerLeuPheHsIlePheAspGIyAspAsn		260
QY	1420	GAAATTTTACAGATCGTGATGATGATTAATCTCAAGATATACTAAATATTTTATGAAAACTA		1475
DB	261	GIuIleuGIuInIleValAspGIuLeuSerGIuAspIleThrIyTyrPheMetLysLeu		280
RESULT 12				
ID	Q25846	PLAFA	PRELIMINARY;	PRT; 280 AA.
AC	Q25846			
DT	01-NOV-1996	(TREMBlrel. 01. Created)		
DT	01-NOV-1996	(TREMBlrel. 01. Last sequence update)		
DT	01-OCT-2002	(TREMBlrel. 22. Last annotation update)		
DE	Liver stage-specific antigen 1 (Fragment).			
GN	Name=LSA-1;			
OS	Plasmodium falciparum.			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5833;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=96065765;	PubMed=747115;	DOI=10.1016/0166-6851(95)00069-D;	
RA	Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,			
RA	Hawley W.A., Collins M.E., Lal A.A.;			
RT	"Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."			
RL	Mol. Biochem. Parasitol. 71:291-294(1995).			
DR	EMBL, L40887; AAC41597.1; -;	Genomic.DNA.		
FT	NON_TER	1		
SO	SEQUENCE	280 AA;	32882 MW;	46763641C81AFC33 CRC64;
Alignment Scores:				
Pred. No.:	1,21e-59	Length:	280	
Score:	1442.00	Matches:	278	
Percent Similarity:	99.64%	Conservative:	1	
Best Local Similarity:	99.29%	Mismatches:	1	
Query Match:	58.00%	Indels:	0	
DB:	2	Gaps:	0	
US-09-837-344-41 (1-1482) x Q25846.PLAFA (1-280)				
QY	640	AGGAGGCTGATACGAAAAAAATTTTGAAGAAAGAAAAAGAAACATGAGATATATTAGCA		699
DB	1	ArgLysAlaAspThrLysLysAsnLeuGIuArgLysLysGIuHsIleGIyAspValLeuAla		20
QY	700	GAGGATTTATATGGTCGTTTGAAGAAATACCAAGCTATAGAACTTCATCAAGAAAAACAAGT		759
DB	21	GIuAspLeuTyrGIyArgLeuGIuIleProAlaIleGIuLeuProSerGIuAsnGIuArg		40
QY	760	GGAATATTATATACACATCATCTTCTTTACCTCAGGACACACAGAGGAATATGATGAGAT		819
DB	41	GIyIyTyrIleProHsIleGInSerSerLeuProGInAspAsnArgGIyAsnSerThrAsp		60

```
QY      820 TCAGAGAAATATCTATATAGAAAAAACAATAGAGATCTATATACAAATGTTGAA 879
      |||
      61 Serlysgluileserilleileglulysrthraenargluserillethrthraenvalglu 80
      |||
QY      880 GCACGAGAGATATATCATTAAGACATCTTGAGAGAAAAAGAGATGTTCAATAAACCA 939
      |||
      81 Glyargargapillehlslysglyhlsleugluglulyslvaspgllyserilleyspro 100
      |||
QY      940 GACCAAAAAAGAGATTAATCTGCTGACATACAAATCTATACCTTAGACAGCTAAATTT 999
      |||
      101 Gluglulysglulvaspserlhaapilleglinaenhsrthleuglulthrvalasnille 120
      |||
QY      1000 TCTGATGTTAATGATTTTCCAAATAAGTAAGTATGAGATGAATAAGTCTGAATATGAC 1059
      |||
      121 Serapvalaenaspheglinileserlystrglulvaspglulileserlleglulysrhap 140
      |||
QY      1060 GATTCATTATATGATGAAGAGAGATGATGAAGCTTAGACGAATTTAAGCTTATGTTG 1119
      |||
      141 AlaSerleuileaspglulglulvaspserglulvaspserglulphelyserroilleval 160
      |||
QY      1120 CAATATGACATTTTCCAGATGAGAGAAAAACATAGAAATTTAATAAGATGAAATTTG 1179
      |||
      161 Glulysrhapasphelglulvaspserlleglyllyrlysglulvaspserllyeu 180
      |||
QY      1180 ATAGAGAAAAATGAAAAATTTAGATGATTTAGTGAAGCAATGAAAAATCATCAGAGAA 1239
      |||
      181 lileglulysrhapserlhaapserlhaapserlhaapserlhaapserlhaapserlhaap 200
      |||
QY      1240 TTAATGAGAAAAATTAATAAGAGAGAAAAATGAAAAACAAAGATTAATAATTT 1299
      |||
      201 Leuserglulglulysllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 220
      |||
QY      1300 AAACCAATGATTAATAAGTTTGTATGATGATGATGATTAATAATAATAATAATAAGTATAGCAG 1359
      |||
      221 Lysproaenaspserlyserleuylrhapglulhlsillellyllyllyllyllyllyllyllyl 240
      |||
QY      1360 GTTAATTAAGAAAAAGAAAAATTCATAAATCATTTGTTTCATATATTTTGACGAGACAT 1419
      |||
      241 Valaenlysglulysglulysrhapelleysrleuphenlsillepheaspglulysrhapen 260
      |||
QY      1420 GAAATTTTACAGATCGTGATGATGATTTATGATGATATTAATAATTTTATGAACTA 1479
      |||
      261 Glulileuglulhlsillevalaspglulvaspserglulvaspserllyrthlyrpheweclyseu 280
      |||
      RESULT 13
      Q25853_PLAFA PRELIMINARY; PRT; 280 AA.
      ID Q25853;
      AC Q25853;
      DT 01-NOV-1996 (TREMBLrel. 01, Created)
      DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
      DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
      DE Liver stage-specific antigen 1 (Fragment).
      GN Name=LSA-1;
      OS Plasmidium falciparum.
      OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmidium.
      RN NCBI_TaxID=5833;
      RP NUCLEOTIDE SEQUENCE.
      RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
      RA Yang C., Shi Y.P., Dohayakumar V., Alpers M.P., Povoa M.M.,
      RA Hawley W.A., Collins W.E., Lal A.A.;
      RT "Sequence variations in the non-repetitive regions of the liver stage-
      RT specific antigen-1 (LSA-1) of Plasmidium falciparum from field
      RT isolates."
      RT Isolates."
      DR EMBL; L40835; AAB59234.1; -; Genomic_DNA.
      FT NON TER 1
      SQ SEQUENCE 280 AA; 32938 MW; 46751C45FADCB33 CRC64;

      Alignment Scores:
      Pred. No.: 2, 08e-59 Length: 280
      Score: 1437.00 Matches: 277
      Percent Similarity: 99.29% Conservative: 1
```

```
Best Local Similarity: 98.93% Mismatches: 2
Query Match: 57.80% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-41 (1-1482) x Q25853_PLAFA (1-280)

QY      640 AGAAGGCTGATTAAGAAAAAATTTAGAAAAAGAAAAAGAAATGAGATATATTACCA 699
      |||
      1 ArgysalaaepThrlyllylvasenleuglulysrlysglulhlsilleglulvaspvalleuala 20
      |||
QY      700 GAGGATTTATATGATGCTGTTTGAAATACACAGCTATAGAACTTCCATCGAAAAATGAAAGT 759
      |||
      21 Glulvaspserlyrlyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllylly 40
      |||
QY      760 GATATTTATATACGACATCAATCTTCTTACCTGACAGCAACAGAGGAGATATGAGAT 819
      |||
      41 Glyllyrlyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 60
      |||
QY      820 TCAGAGAAATATCTATATAGAAAAAACAATAGAGATCTATATACAAATGTTGAA 879
      |||
      61 Serlysgluileserilleileglulysrthraenargluserillethrthraenvalglu 80
      |||
QY      880 GCACGAGAGATATATCATTAAGACATCTTGAGAGAAAAAGAGATGTTCAATAAACCA 939
      |||
      81 Glyargargapillehlslysglyhlsleugluglulyslvaspgllyserilleyspro 100
      |||
QY      940 GACCAAAAAAGAGATTAATCTGCTGACATACAAATCTATACCTTAGACAGCTAAATTT 999
      |||
      101 Gluglulysglulvaspserlhaapserlhaapserlhaapserlhaapserlhaapserlhaap 120
      |||
QY      1000 TCTGATGTTAATGATTTTCCAAATAAGTAAGTATGAGATGAATAAGTCTGAATATGAC 1059
      |||
      121 Serapvalaenaspheglinileserlystrglulvaspglulileserlleglulysrhap 140
      |||
QY      1060 GATTCATTATATGATGAAGAGAGATGATGAAGCTTAGACGAATTTAAGCTTATGTTG 1119
      |||
      141 Asperleuileaspglulglulvaspserlhaapserlhaapserlhaapserlhaapserlhaap 160
      |||
QY      1120 CAATATGACATTTTCCAGATGAGAGAAAAACATAGAAATTTTATAAGATGAAATTTG 1179
      |||
      161 Glulysrhapasphelglulvaspserlleglyllyrlyllyllyllyllyllyllyllyllyl 180
      |||
QY      1180 ATAGAGAAAAATGAAAAATTTAGATGATTTAGTGAAGCAATGAAAAATCATCAGAGAA 1239
      |||
      181 lileglulysrhapserlhaapserlhaapserlhaapserlhaapserlhaapserlhaapserl 200
      |||
QY      1240 TTAATGAGAAAAATTAATAAGAGAGAAAAATGAAAAACAAAGATTAATAATTT 1299
      |||
      201 Leuserglulglulysllyllyllyllyllyllyllyllyllyllyllyllyllyllyllylly 220
      |||
QY      1300 AAACCAATGATTAATAAGTTTGTATGATGATGATTAATAATAATAATAATAAGTATAGCAG 1359
      |||
      221 Lysproaenaspserlyserleuylrhapglulhlsillellyllyllyllyllyllyllyllyl 240
      |||
QY      1360 GTTAATTAAGAAAAAGAAAAATTCATAAATCATTTGTTTCATATATTTTGACGAGACAT 1419
      |||
      241 Valaenlysglulysglulysrhapelleysrleuphenlsillepheaspglulysrhapen 260
      |||
QY      1420 GAAATTTTACAGATCGTGATGATGATTTATGATGATATTAATAATTTTATGAACTA 1479
      |||
      261 Glulileuglulhlsillevalaspglulvaspserglulvaspserllyrthlyrpheweclyseu 280
      |||
      RESULT 14
      Q25854_PLAFA PRELIMINARY; PRT; 280 AA.
      ID Q25854;
      AC Q25854;
      DT 01-NOV-1996 (TREMBLrel. 01, Created)
      DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
      DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
      DE Liver stage-specific antigen 1 (Fragment).
      GN Name=LSA-1;
      OS Plasmidium falciparum.
      OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmidium.
```

NCBI_TaxID=5833;
[1]
NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Uchayakumar V., Alpers M.P., Pova M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL/ L40836; AAC41604.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32566 MW; 61851C5AD6E1A211 CRC64;
Alignment Scores:
Pred. No.: 2,86e-59 Length: 280
Score: 1434.00 Matches: 276
Percent Similarity: 99.29% Conservative: 2
Best Local Similarity: 98.57% Mismatches: 2
Query Match: 57.68% Indels: 0
Gaps: 0
US-09-837-344-41 (1-1482) x Q25854_PLAFA (1-280)
QY 640 AGGAAGCTGTATCGAAAAAATTTAGAAAGAAAAAGAACATGAGATATTTAGCA 699
DB 1 Arglysaiaaaprrlrlvlyasnlleuqlurgylysglulhsglyasrvalleualla 20
QY 700 GAGGATTTATATGCGCTTTAGAAAATACAGCTATAGAACTTCATCAGAAATGAACGT 759
DB 21 Gluaspneuylrglyasglleuqlulleprovalilleglulneuprosesgluasnlgularg 40
QY 760 GGATATTATATACACATCAATCTTCTTACCTCAGACAACAGAGGAATAGTAGAT 819
DB 41 Glylyrlyrllleprohlsnglnserleuproglnasrpaanrgglyasnserargasp 60
QY 820 TCCAGAGAAATATCTAATATGAAAAAACAATAGAGAACTTATTTACAACTTAATTTGAA 879
DB 61 Serlysglulleserillelleglulvsthasnrglulserilleththrasnvalglu 80
QY 880 GGACGAAGGATATATCATTAAGACATCTTGAAGAAAAGAAAGATGGTTCAATAAACA 939
DB 81 Glylrglrgaaprllehlslvlyhlsbeuqlulvlyslvshapgllyserillelypro 100
QY 940 GAACAAAAGAGATTAATCTGCTGACATACAAAATCATACATTAGAGACATTAATTT 999
DB 101 Glulnlysgluasrlyserlaseprilleglinsnhtsthrleuqluthrvalasnille 120
QY 1000 TCTGATGTTAATGATTTTCAAAATAGTAGAGATGAATTAAGCTCGAATATGAC 1059
DB 121 Serasrvalasnarpheglilleserlystrgluaspgrlulleserlleglulyrasp 140
QY 1060 GATTCATTAATAGATGAAGAAAGAAAGATGAGACTTAGAGCAATTTAGCTATTGTTG 1119
DB 141 Aspsertleullleasrpglulglulnspasrpglulnspasrpglulnspasrpglulnsp 160
QY 1120 CAATATGACAAATTTCCAAAGATGAAGAAACAATAGCAATTTTAAGAAGCTAGAAATTTG 1179
DB 161 Glnlyrgrasrpaenpheglinsrpglulnlelleglyllyrlysglulnleuqlunspreu 180
QY 1180 ATAGAGAAAAATTTAGATGATTTAGATGAAGAAATAGAAATTCATCAGAGAA 1239
DB 181 lileglulvbasnlgulnleuasnlpasrpglulglulleglulvsersegluln 200
QY 1240 TTATGTGAAGAAAAATATAAAAAAGAAAGAAATAGAAAAACAAGATTAATATTTT 1299
DB 201 Leuserrglulnlgulargllylelyserglulvlystrglulvsthrlyasrpaenphe 220
QY 1300 AAACCAAAATGATTAAGATTTGTATGATGACATATTAATAAATATTAATAATGATTAAGAC 1359
DB 221 Lysrproasnarpleserleuylrshrglulnsllelyslvlylyasrvallysgln 240
QY 1360 GTTAATAAGAAAGAAAAATTCATAAAATCATTTGTTTCATATATTTGACGAGACAAAT 1419

DB 241 Valasnlysglulvsglulvlysrphellelyserleuphenlilelphasrpglyasrpaen 260
QY 1420 GAAATTTTACAGATCGTGATGATGATTAATCTGAAGATTAATTAATTTTATGAACATA 1479
DB 261 Glulileuqlulnlevalaasrpglulnleuaserglulnasprrlethrlyrphemetylslenu 280
RESULT 15
Q25845_PLAFA PRELIMINARY; PRT; 280 AA.
ID Q25845_PLAFA
AC Q25845;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Uchayakumar V., Alpers M.P., Pova M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL/ L40886; AAB59232.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32908 MW; FDEB29F34DCAT7B9 CRC64;
Alignment Scores:
Pred. No.: 3.19e-59 Length: 280
Score: 1433.00 Matches: 276
Percent Similarity: 98.93% Conservative: 1
Best Local Similarity: 98.57% Mismatches: 3
Query Match: 57.64% Indels: 0
Gaps: 0
US-09-837-344-41 (1-1482) x Q25845_PLAFA (1-280)
QY 640 AGGAAGCTGTATCGAAAAAATTTAGAAAGAAAAAGAACATGAGATATTTAGCA 699
DB 1 Arglysaiaaaprrlrlvlyasnlleuqlurgylysglulhsglyasrvalleualla 20
QY 700 GAGGATTTATATGCGCTTTAGAAAATACAGCTATAGAACTTCATCAGAAATGAACGT 759
DB 21 Gluaspneuylrglyasglleuqlulleprovalilleglulneuprosesgluasnlgularg 40
QY 760 GGATATTATATACACATCAATCTTCTTACCTCAGACAACAGAGGAATAGTAGAT 819
DB 41 Glylyrlyrllleprohlsnglnserleuproglnasrpaanrgglyasnserargasp 60
QY 820 TCCAGAGAAATATCTAATATGAAAAAACAATAGAGAACTTATTTACAACTTAATTTGAA 879
DB 61 Serlysglulleserillelleglulvsthasnrglulserilleththrasnvalglu 80
QY 880 GGACGAAGGATATATCATTAAGACATCTTGAAGAAAAGAAAGATGGTTCAATAAACA 939
DB 81 Glylrglrgaaprllehlslvlyhlsbeuqlulvlyslvshapgllyserillelypro 100
QY 940 GAACAAAAGAGATTAATCTGCTGACATACAAAATCATACATTAGAGACATTAATTT 999
DB 101 Glulnlysgluasrlyserlaseprilleglinsnhtsthrleuqluthrvalasnille 120
QY 1000 TCTGATGTTAATGATTTTCAAAATAGTAGAGATGAATTAAGCTCGAATATGAC 1059
DB 121 Serasrvalasnarpheglilleserlystrglulnspasrpglulnspasrpglulnspasrpglulnsp 140
QY 1240 TTATGTGAAGAAAAATATAAAAAAGAAAGAAATAGAAAAACAAGATTAATATTTT 1299
DB 141 Aspsertleullleasrpglulglulnspasrpglulnspasrpglulnspasrpglulnsp 160

```

Qy 1120 CAATATGACAAATTTCCAGATGAGAAAAATATGAAATTTATTAAGAACTAGAGATTG 1179
    |||
Db 161 GlnTyrAspAsnPhcGlnAspGlnuAsnIleGlyIleTyrLysGlnuLeuGlnuAspLeu 180
    |||
Qy 1180 ATAGAGAAAAATGAAATTTAGATGATTTAGATGAAAGAAATGAAAAATCATCAGAGAA 1239
    |||
Db 181 ILeGlnLysAsnGlnuAsnLeuAspLeuAspGlnuGlyIleGlnLysSerSerGlnu 200
    |||
Qy 1240 TTATCTGAAGAAAAATTAATAAGAAAGAAATATGAAAAACAAAGATATTAATTTT 1299
    |||
Db 201 LeuSerGlnuGlnuLysIleLysLysGlnuLysTyrGlnuLysThrLysAspAsnAspPhe 220
    |||
Qy 1300 AAACCAATGATTAAGATTTGTATGATGATGATTAATAATAATAATAATAATAATAAGCAG 1359
    |||
Db 221 LysProAsnAspLysSerLeuTyrAspGlnuHisIleLysLysTyrLysAsnValLysGln 240
    |||
Qy 1360 GTTAATTAAGAAAGAAAAATTCATTAATAATCATTTGTTTCATATATTTTGACGAGACAAT 1419
    |||
Db 241 ValAsnLysGlnuLysGlnuLysPheIleLysSerLeuPheHisIlePheAspGlnuAspAsn 260
    |||
Qy 1420 GAAATTTTACAGATCGTGATGATGATTAATCTGAAGATATTAATAATATTTTATGAAACTA 1479
    |||
Db 261 GlnIleLeuGlnIleValAspGlnuLeuSerGlnuAspIleThrLysTyrPheMetLysLeu 280
    |||

```

Search completed: December 7, 2005, 06:55:28
 Job time : 221.5 secs

XX MPI: 1992-299985/36.
DR N-PSDB; AAQ28119.
XX
PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
PS Claim 2: Fig 8-10; 81pp; French.
XX
CC The 3' part of the P. falciparum liver-stage specific antigen (LSA) gene
CC codes for a polypeptide sequence which carries a T cell epitope
CC characteristic of a protein produced in hepatocytes infected with
CC P. falciparum. The polypeptide can be used in the preparation of vaccines
CC against malaria. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 493 AA;

Alignment Scores:

Pred. No.:	2, 81e-177	Length:	493
Score:	2478.00	Matches:	493
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.68%	Indels:	0
DB:	2	Gaps:	0

US-09-837-344-41 (1-1482) x AAR26944 (1-493)

QY 1 CAAAGCAACAAAGCGATTTAGAACAGAGACGTGCTAAAGAAAAGTTGCAAGACAA 60
DB 1 GInGInGInGInSeraspLeuGInGInGInGInArgAlaLysGInLysLeuGInGInGIn 20
QY 61 CAAAGCGATTTAGAACAAAGTACTTGCTAAAGAAAAGTTGCAAGAGCAAGCGAT 120
DB 21 GInSeraspLeuGInGInSeraspLeuAlaLysGInLysLeuGInGInGInSerasp 40
QY 121 TTAGAACAGAGACGCTTCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAA 180
DB 41 LeuGInGInGInLysLeuAlaLysGInLysLeuGInGInGInSeraspLeuGInGIn 60
QY 181 GAGAGACGTGCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAGACGT 240
DB 61 GInArgArgAlaLysGInLysLeuGInGInGInSeraspLeuGInGInGInArg 80
QY 241 GCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAGACGTGCTAAAGAA 300
DB 81 AlaLysGInLysLeuGInGInGInSeraspLeuGInGInSeraspLeuAlaLysGIn 100
QY 301 AAGTTCAAGAGACGCAAGCGATTTAGAACAGAGACGTGCTAAAGAAAAGTTGCA 360
DB 101 LysLeuGInGInGInGInSeraspLeuGInGInGInArgArgAlaLysGInLysLeuGIn 120
QY 361 GAAACAACAAAGCGATTTAGAACAGAGACGTGCTAAAGAAAAGTTGCAAGACAA 420
DB 121 GInGInGInSeraspLeuGInGInGInGInArgArgAlaLysGInLysLeuGInGInGIn 140
QY 421 AGCGATTTAGAACAGAGACGCTTCTAAAGAAAAGTTGCAAGACAAAGCGATTTA 480
DB 141 SeraspLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 160
QY 481 GAAACAAGAGACGTGCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAG 540
DB 161 GInGInGInLysArgArgAlaLysGInLysLeuGInGInGInGInSeraspLeuGInGIn 180
QY 541 AGACGTGCTAAAGAAAAGTTGCAAGACAAAGCGATTTAGAACAGAGACGTGCT 600
DB 181 ArgArgAlaLysGInLysLeuGInGInGInGInGInGInGInGInGInGInGInGInGIn 200
QY 601 AAAAGAAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGAAAGCGCTGATAGAAAA 660
DB 201 LysGInLysLeuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 220
QY 661 AATTGAAAAGAAAAGAAATGAGATATATTAGCAGAGATTTATGTCGTTTA 720

DB 221 AsnLeuGInArgLysLysGInLysGInLysAlaSerPheLeuAlaGInAspLeuTyGInArgLeu 240
QY 721 GAAATACCAAGCTATAGAACTTCCATCAGAAAATAGAACTGGATATTATATACCATCAA 780
DB 241 GInLysPheAlaLysGInLysPheSerGInLysGInLysGInLysGInLysGInLysGIn 260
QY 781 TCTTCTTACCTCAGAGCAACAGAGGAATAGTAGATTTCCAGAAAATATCTTAATA 840
DB 261 SerSerLeuProGInAspAsnArgLysSerArgAspSerLysGInLysSerLysLys 280
QY 841 GAAAAAACAATTAGAGATCTTATTCACAAATGTTGAAAGACGAAAGGATTTACATAA 900
DB 281 GInLysSerAsnArgGInLysSerLysLysLysLysLysLysLysLysLysLysLysLys 300
QY 901 GGACATCTTGAAGAAAAGAAAGATGCTCAATTAACCAAGAACAAAGAAAGATTAATCT 960
DB 301 GInLysLeuGInLysLysLysPheLysPheLysLysLysLysLysLysLysLysLysLys 320
QY 961 GCTGACATACAAATCATATACATTTAGAGACAGTAAATATTTCTGATGTTAATGATTTCAA 1020
DB 321 AlaAspPheGInAsnHisThrLeuGInThrValAlaSerLysSerValAsnAspPheGIn 340
QY 1021 ATTAAGTAAATAGAGATGAAATTAAGTCTGATTTGACGATTTCTTAATAGATGAAGA 1080
DB 341 LysSerLysTyArgLysAspGInLysSerAlaGInTyArgAspSerLeuLysAspGIn 360
QY 1081 GAAGATGATGAAGATTTAGACGAATTTAAGCTTATGCAATATGACATTTCCAGAT 1140
DB 361 GInAspAspGInAspLeuAspLysPheLysPheLysPheLysPheLysPheLysPheLys 380
QY 1141 GAAAGAAAATAGAAATTTATTAAGAACTAGAGATTTGATAGAGAAAATGAATAATTA 1200
DB 381 GInGInGInLysGInLysGInLysGInLysGInLysGInLysGInLysGInLysGIn 400
QY 1201 GATGATTTAGATGAAGAAATAGAAAATCATAGAGAAATTTCTGAAGAAAATTAATA 1260
DB 401 AspAspLeuAspGInGInLysGInLysSerGInLysSerGInLysSerGInLysLys 420
QY 1261 AAAGGAAAATAATGAAGAAAACAAAGGATAATTTTAACCAATGATTAAGTTTG 1320
DB 421 LysGInLysLysTyGInLysThrLysAspAsnAsnPheLysPheAsnAspLysSerLeu 440
QY 1321 TATGATGACATATTAATAAATAATTAATAATGAAGCAAGTTAATAGGAAAAGAAAA 1380
DB 441 TyrAspGInLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 460
QY 1381 TTCATTAATCATTTGTTTCATATATTGACGAGACATGAATTTTACAGATCGTGAT 1440
DB 461 PheLysSerLeuPheHisLysLysPheAspGInLysAspAsnGInLysLeuGInLysLys 480
QY 1441 GAGTTATCGAGATATTAATTAATTTTATGAAGACTA 1479
DB 481 GInLeuSerGInAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 493
RESULT 2
ID ADO21942 standard; protein; 456 AA.
AC ADO21942;
XX 12-AUG-2004 (first entry)
XX
XX LSA-NRC(H) construct protein derived from Malaria parasite LSA-1.
XX
XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
XX parasite; vacuole; antimalarial; vaccine; immunotherapy;
XX malaria parasite P. falciparum; LSA-NRC(H) construct.
OS Plasmodium falciparum.
XX Synthetic.
XX
XX MO2004044167-A2.


```
Db      428 GlnIleValAspGluArgLeuSerGlnAspIleThrIleTyrPheMetLeu 445
|||||
RESULT 4
AAR26941
ID      AAR26941 standard; protein; 316 AA.
XX
AC      AAR26941;
XX
DT      25-MAR-2003 (revised)
DT      08-FEB-1993 (first entry)
XX
DE      P.falciparum LSA-R-NR protein.
XX
KM      Malaria; hepatocyte; sporozoite; plasmod DG 536; T-cell epitope;
XX      paludism; liver stage-specific antigen.
XX
OS      Plasmodium falciparum.
XX
FH      Key
FH      Region
FH      Location/Qualifiers
FT      1..209
FT      /label= repeat region
FT      /note= "contains 12 x 17mer repeats"
FT      210..316
FT      /label= non-repeat_region
XX
XX      MO9213884-A1.
XX
XX      20-AUG-1992.
XX
XX      05-FEB-1992; 92MO-FR000104.
XX
XX      05-FEB-1991; 91FR-00001286.
XX
XX      (INSP ) INST PASTEUR.
XX
XX      Guerimarchand C, Druilh P;
XX
XX      WPI; 1992-299985/36.
XX      DR      N-PSDB; AA028115.
XX
XX      Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX      PT      vaccination against, treatment of and diagnosis of malaria.
XX
XX      PS      Disclosure; Fig 1; 81pp; French.
XX
XX      A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
XX      CC      gIII was used to transform E.coli. The expression library was screened
XX      CC      with human antisera against antigens of all stages of P. falciparum. The
XX      CC      library was rescreened with antibodies affinity- purified on a clone
XX      CC      which was able to recognise antibodies specific to the hepatic phase.
XX      CC      About 40 clones were detected which produced a characteristic LSA
XX      CC      epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
XX      CC      containing a 12-repeat region followed by a non-repeat region. Preferred
XX      CC      antigenic polypeptides of the invention are derived from the amino acid
XX      CC      sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX      SQ      Sequence 316 AA;

Alignment Scores:
Pred. No.:      1,356-108      Length:      316
Score:          1561.00      Matches:      316
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    62.79%      Indels:      0
DB:             2      Gaps:      0

US-09-837-344-41 (1-1482) x AAR26941 (1-316)
QY      13 AGCGATCTGACACAGAGACGTGCTAAAGAAAGTTGCAAGACAAACAAAGCATTTA 72
DB      1 SerAspLeuGluGlnGluArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeu 20
```

```
QY      73 GAACAAGATAGACTTGCTAAAGAAAGTTACAGACAGCAAGCGATTTAGAACAGAG 132
DB      21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGln 40
QY      133 AGACTTGCTAAAGAAAGTTGCAAGACAAACAAAGCGATCTAGAACAAAGAGACGTGCT 192
DB      41 ArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgAla 60
QY      193 AAAGAAAGAGTTGCAAGACAAACAAAGCGATTTAGAACAGAGAGACGTGCTAAAGAAAG 252
DB      61 LysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLysGluLys 80
QY      253 TTGCAAGACAAACAAAGCGATTTAGAACAGATAGACTTGCTAAAGAAAGTTACAGAG 312
DB      81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGln 100
QY      313 CAGCAAAAGCGATTTAGAACAGAGACGTGCTAAAGAAAGTTGCAAGACAAACAAAGC 372
DB      101 GlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnSer 120
QY      373 GATTTAGAACAAAGAGACGTGCTAAAGAAAGTTGCAAGACAAACAAAGCGATTTAGAA 432
DB      121 AspLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln 140
QY      433 CAAGAGACTTGCTAAAGAAAGTTGCAAGACAAACAAAGCGATTTAGAACAGAGAGAG 492
DB      141 GlnGlnArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArg 160
QY      493 CGTGCTAAAGAAAGTTGCAAGACAAACAAAGCGATTTAGAACAGAGAGACGTGCTAA 552
DB      161 ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLys 180
QY      553 GAAAGTTGCAAGACAAACAAAGCGATTTGCAAGAGAGACGTGCTAAAGAAAGTTG 612
DB      181 GluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLysGluLysLeu 200
QY      613 CAAGAGACGCAAGAGATTAGAACAAAGAGAGCGATTCAGAAAAAAATTTAGAAAGA 672
DB      201 GlnGlnGlnGlnGlnArgAspLeuGlnGlnGlnGlnArgAlaAspThrIleLysSerAsnLeuGlnArg 220
QY      673 AAAAGAGACATGAGATATATTAGACAGAGATTTATGTCGTTTGAATPCCAGCT 732
DB      221 LysLysGlnGlnGlnGlnAspIleLeuAlaGlnAspLeuTyrGlnGlnGlnIleProAla 240
QY      733 ATAGAACTTCATCAGAAATGAACGTGATATTATTATACCAATCAATCTCTTTACCT 792
DB      241 IleGlnLeuProSerGlnGlnGlnGlnArgGlyTyrTyrIleProHisGlnSerSerLeuPro 260
QY      793 CAGAGACAAACAGAGGAGATTAGAGATTTCCAGAAATATCTATATAGAAAAACAAT 852
DB      261 GlnAspAsnArgGlyAsnSerArgAspSerLysGlnIleSerIleIleGlnLysThrAsn 280
QY      853 AGAGAACTCTATTACAAATGTTGAGAGCAAGAGGATATACATAAAGACATCTTGAA 912
DB      281 ArgGlnSerIleThrThrAsnValGlnGlyArgArgAspIleHisLysGlnHisLeuGln 300
QY      913 GAAAAAGAAAGATGTTCAATPAAACCAAGACAAAGAAAGAAATATCT 960
DB      301 GluLysLysAspGlySerIleLysProGlnGlnGlnGlnLysAspLysSer 316

RESULT 5
AAR05766
ID      AAR05766 standard; protein; 462 AA.
XX
AC      AAR05766;
XX
DT      25-MAR-2003 (revised)
DT      05-NOV-1990 (first entry)
XX
DE      Portion of peptide antigen to malarial sporozoite.
XX
XX      Malaria; sporozoite; vaccine; exoerythrocytic parasites; tetanus toxoid.
XX
```

Alignment Scores:	
Pred. No.:	2 94e-64
Score:	69.50
Percent Similarity:	61.88
Best Local Similarity:	49.68
Query Match:	39.00*
DB:	2
Length:	462
Matches:	232
Conservative:	57
Mismatches:	95
Indels:	83
Gaps:	10

```

QY      1  CAAGAACCAACAAAGCATCTAGAACCAAGAGAGACGTCCTAAAGAAAAGTTGCACGAACAA  60
Db      74  GlnGlyGlnGlnSerAspLeuGlnGlnGluValAlaIlySglIyIlyLeuGlnGlnGln  93
QY      61  CAAAGCGATTTAGAACCAAGATAGACTTGTCTAAAGAAAAGTTACAGAGCAGCAAAAGCAT  120
Db      94  GlnSerAspLeuGlnGlnGlnGlnGluValAlaIlySglIyIlyLeuGlnGlnGlnGlnSerAsp  113
QY      121  TTAGAACCAAGAGAGACTTGCTTAAGAAAAGTTGCCAAGAACCAAAAGCGATCTTAGAACAA  180
Db      114  LeuGlnGlnAspArgLeuAlaIlySglIyIlyLeuGlnGlnGlnGlnSerAspLeuGlnGln  133
QY      181  GAGAGACGTCCTAAAGAAAAGTTGCCAAGAACCAAAAGCGATTTAGAACCAAGAGAGACGT  240
Db      134  GluArgArgAlaIlySglIyIlyLeuGlnGlnGlnGlnSerAspLeuGluArgThrIlyValAla  153
QY      241  GCTAAAGAAAAGTTGCACAGAACCAAGCGATTTTAGAACCAAGATAGACTTGCTTAAGAA  300
Db      154  SerThrGluThrLeuHisGlnGlnGlnGlnSerAspLeuGlnGlnGlnGluValAlaIlySglI  173
QY      301  AAGTTACAAAGAGCAGCAAAAGCGATTTTAGAACCAAGAGAGACGTGCTTAAGAAAAGTTGCA  360
Db      174  IlyLeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGlnGluValAlaIlySglIyIlyLeuGln  193
QY      361  GAACCAACAAAGCGATTTAGAACCAAGAGAGACGTGCTTAAGAAAAGTTGCACGAACCAAA  420
Db      194  GlnGlnGlnSerAspLeuGlnGlnGlnGluValAlaIlySglIyIlyLeuGlnGlnGlnGln  213
QY      421  ACGCATTTTAGAACCAAGAGACTTGTCTAAAGAAAAGTTGCACGAACCAAGCGATTTTA  480
Db      214  SerAspLeuGlnGlnGlnGlnGluArgAlaIlySglIyIlyLeuGlnGlnGlnGlnSerAspLeu  233
QY      481  GAACCAAGAGAGACGTGCTTAAGAAAAGTTGCACGAACCAAGCGATTTAGAACCAAGAG  540
Db      233  SerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln  252

```

RESULT 6	
AA26943	
ID	AA26943 standard; protein, 318 AA.
XX	
AC	AA26943;

DT	25-MAR-2003	(revised)
DT	08-FEB-1993	(first entry)
XX		
DE	P.falciparum LSA N-terminal portion.	
XX		

US-09-837-344-41 (1-1482) x AAB26937 (1-117)

QY 610 TTGCAAGACGACCAAGAGATTAGAACAAAGAAAGCTGATACGAAAAAATTTAGAA 669

DB 1 LeuGInGluGInGlnAhrGAspLeuGInGlnAhrGlySAIAAspThrLysLysAsnLeuGIn 20

QY 670 AGAAAAAGGACATGAGAGATATATTAGACAGAGATTTATATGTCGTTTAAAGATACCA 729

DB 21 AAGLysLysGluHISGlyAspLLeuAlaGluAspLeuTyrGlyArgLeuGluLeu 40

QY 730 GCTATGAACTTCATCAGAAATGAAACGTGATATTATATACCAATCAATCTCTTTA 789

DB 41 AlaIleGluLeuProSerLeuAsnGlnAhrGlyTyrTyrIleProHISGlnSerLeu 60

QY 790 CCTCAGGACACAGAGGAAATGTAGAGATTTCCAAAGAAATTCATTAATAGAAAAACA 849

DB 61 ProGlnAspAsnAhrGlyAsnSerAhrGAspSerLysGluIleSerIleIleGluLysThr 80

QY 850 AATAGAGATCTATTACACAAATGTGAAGACGAGAGGATATCATTAAGACATCTT 909

DB 81 AAsnAhrGluSerIleThrThrsAsnValGluGlyAhrGAspLLeuHISLysGlyHISLeu 100

QY 910 GAAGAAAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAAGATTAATCT 960

DB 101 GluGluLysLysAspLysSerIleLysProGluGlnLysGluAspLysSer 117

RESULT 8

ABG17301

ID ABG17301 standard; protein; 419 AA.

XX ABG17301;

AC 18-FEB-2002 (first entry)

XX 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #17292.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

XX 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YF,

PI MPI; 2001-639362/73.

XX N-PSDB; AAS81468.

DR New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 47660; 103bp; English.

PS The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

XX Sequence 419 AA;

SO

Alignment Scores:

Pred. No.: 7,81e-20 Length: 419

Score: 376.50 Matches: 149

Percent Similarity: 44.34% Conservative: 35

Best Local Similarity: 35.90% Mismatches: 76

Query Match: 14.26% Indels: 155

De: 4 Gaps: 25

US-09-837-344-41 (1-1482) x ABG17301 (1-419)

QY 1217 CCTCATCAATCAATCAATTAATTTCTCTATCAATCTTCTAGTTCTTATTA 1158

DB 136 ProAlaProAlaSerSerProCysSer--ArgSerLeuSerSerAhrSerLeu-- 153

QY 1157 ATTCTATGTTTCTTCATCTTGAA-----TTGCTATATTGCA 1116

DB 154 -----SerSerSerCysAhrGValSerCysCysHISerLeuSerSerCysCys 169

QY 1115 ATAGCTTAATTCCTTAAGTCTTCAATCATCTTCTTCTCATCTTAATTAATCAATCGTCA 1056

DB 170 ArgSerLeuSerCysSerAhrSerSerSerThrSerAhrGSerSerAhrSerSer 189

QY 1055 TATTGAGACATTATTCATCCCATCACTTACTTATTGAAATCAATTAACATCAGAAATA 996

DB 190 SerSerLeuSerSerSerSerLeu----- 198

QY 995 TTTACTGCTCAATAGTATGATTTGTATGTCAGAGATTATCTTTTGTCTGTGCT 936

DB 199 -----SerCysSerSerCysAhrGLeuSer-----CysSer----- 208

QY 935 TTTATTAACCATCTTCTTCTTCTCAAGATGCTTATATGTAATCCCTTGCTCTTCA 876

DB 208 ----- 208

QY 875 ACATTGTGTAATAGATTCTCTATTGTTTTTCTATTATAGATTATTCCTTGAAATCT 816

DB 208 ----- 208

QY 815 CTACTATTCCCTGTGTGCTGAGTAAGAAGATTGTGTATATATATATCCACGT 756

DB 209 -----Asn 209

QY 755 TCATTTTGTGATGAGATTCTATAGCTGATTTCTTAACGACATATTAATCTCTGTCT 696

DB 210 SerPheSerThrSerSerSerLeuSerLeuSer----- 221

QY 695 AATATATATCCCATGTTCTTTTCTTCTAATATTTTTCGTATACAGCTTCTCTTGT 636

DB 222 -----SerSerLeuSerCys 226

QY 635 TCTAATATCTTTGTGCTCTGCAACTTTTCTTAGACAGCTCTCTGTTCTTAATG 576

DB 227 Ser---SerCysCysLeuSerCysSerAhrSerPheSerThrSerSerSerLeuSer 245

QY 575 CTGTGTGTTCTTGAACATTCTTTAGACAGCTCTCTGTTCTTAATGCTGTTGTCT 516

DB 246 LeuSerSerSerLeuSerCysSerSerCysAhrGLeuSerCysSerAhrSerPhe----- 263


```
QY 673 AAAAAGCAATGAGATATATTAGCAGATTTATATGTCGTF----- 717
Db 794 LySerGluArgGly-----ArgGlnArgLeuSerAlaArgProProLeuArgGlu 810
QY 718 -----TTGAAATACACAGCTTATAGAACTTCATCAGAAAAAGAACTGGATTTAT 768
Db 811 GlnArgGluArgGlnLeuArgAlaGlnGlnArgGlnGlnArgGlnGlnArg-----Phe 828
QY 769 ATACCAATCAATCTTTCTTTACCTGAGCAACAGAGGGAATGTAAGATTCCAGGAA 828
Db 829 LeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 848
QY 829 ATATCTATATGAAAAAACAATAGA-----GAATCTATTACA 867
Db 849 LeuGlnPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 868
QY 868 ACAATGTTGAAGA-----CGAAGGATATATACATMAAGACATCTTGAGAAAAAGAAAGT 924
Db 869 GlnGlnGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 888
QY 925 GGTCAATMAAACCAAGAACAAAAAGATTAATCTGCTGACATCAAAATCATCATTA 984
Db 889 GlnLysTrpArgTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 906
QY 965 GAGACAGTAAATTTTCATGTTATGATTTTCAAAATAGTAAGTAAGAGATGAAATA 1044
Db 907 -----TyrAlaLysProAlaLeuGlnGlnGlnGlnGln 916
QY 1045 AGTGCATATATGACGATTCATTAATATAGTGAAGAAGATGATGAAGATTAAGCA 1104
Db 917 ArgLysGln-----GlnGlnLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 934
QY 1105 TTTAAGCTATTGTGCATATGACAAATTTCCAAATGAAAGAAACATGGAATTTATTA 1164
Db 935 -----ArgGlnLysArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 948
QY 1165 GAACCTAGAAAGATTTGATAGAGAAAAATTAAGATTTGATGAGCAAGAAATGAGA 1224
Db 949 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 966
QY 1225 AATATCATCAGAAATTTATCTGAGAAAAATTAATAAGAAAGAAATATGAAAAACA 1284
Db 967 LysArgArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 986
QY 1285 AAGATTAATTAATTTT---AAACCAATGATTAAGTTGATGATGAGCATTTAAAA 1341
Db 987 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1006
QY 1342 TTTAAAAATGATAGCAAGTTAATAAGAAAAAGAAAAATTCATAAAA 1389
Db 1007 TyrArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1022
RESULT 10
ADD48869
ID ADD48869 standard; protein; 1898 AA.
XX
XX ADD48869;
XX
XX 02-DEC-2004 (revised)
XX 29-JAN-2004 (first entry)
XX
XX Human Protein AAA65582, SEQ ID NO 14580.
XX
XX Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNr; Chung.
XX
XX Homo sapiens.
XX Unidentified.
XX
XX MO2003016475-AZ.
XX
```

```
PD 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GENO) GEN HOSPITAL CORP.
XX
XX (FARO) BAYER AG.
XX
XX Woolf C, D'Urso D, Befort K, Costigan M;
XX
XX MPI; 2003-268312/26.
XX
XX GENBANK; AAA65582.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (SNr), chronic constriction
XX injury (CCI) and spared nerve injury (SNr)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (described in Table 3
XX of the specification) which is differentially expressed during pain.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1898 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1,02e-19 Length: 1898
XX Score: 376.00 Matches: 137
XX Percent Similarity: 47.39% Conservative: 117
XX Best Local Similarity: 25.56% Mismatches: 174
XX Query Match: 15.12% Indels: 108
XX DB: 7 Gaps: 24
XX
XX US-09-837-344-41 (1-1482) x ADD48869 (1-1898)
QY 1 CAAGAACAAACAAAGCATCTGAACAGAGAGCGTGAAGAAAGTTGCAAGACAA 60
Db 522 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 541
QY 61 CAAGCGATTTAGAACAAAGATGACTTGCTAAAGAAAGTTACAAAGACGAGCAAGCGAT 120
Db 542 Gln-----GlnGlnGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 558
QY 121 TTGAACAA---GAGACACTTGCTTAAGAAAAATTGCAAGAAACAAAGCATCTAGAA 177
Db 559 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 578
```

QY	178	CAGAGAGACGCTCTAAAGAAAAGTTGCAAGACAAACAAGC-----	219
DB	579	GINGLIATGARGXGLUGLInARGLLeuLYbRRGInGLInGLIuARGARGbPGLInLeu	598
QY	220	GATTTAGAACAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAAGGATTTAGAA	279
DB	599	LYbARGInGLUGLInARGbGLInGLInARGLLeuLYbARGbGLInGLInGLInARGbGLIn	618
QY	280	CAGATAGACCTTGCTTAAAGAAAAGTTTACAAAGACGCAAGCGATTTTAGAACAGAGCA	339
DB	619	GIN---ARGLeuLYbARGbGLUGLInuVALGLInARG-----LeuGLInGLInGLIn	633
QY	340	CGTCTTAAAGAAAAGTTG-----CAAGAACAAAGCGATTTA---	378
DB	634	ARGARGbPGLInuARGLeuLYbARGbGLUGLInuPROGLInGLInuARGbGLInLeu	653
QY	379	-----GAACAAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA	429
DB	654	LYbSERGLUGLInGLInGLInuARGbGLInuSERGLInLeuARGbGLInGLInGLInuARG	673
QY	430	GAACAAGAGACGCTCTTAAAGAAAAGTTGCAAGAACAA-----CAAGAC	474
DB	674	ARGbGLInuARGbLeuLYbARGbGLUGLInuGLInuARGbLeuGLInuARGbLeuLYbARG	693
QY	475	GATTTAGAACAAAGAGACGCTCTTAAAGAAAAGTTGCAAGAACAAACAAGC-----	525
DB	694	GINHbGLInGLInGLInuARGbGLInGLInuLEuVALGLInGLInGLInGLInuVALARG	713
QY	526	-----GATTTAGAACAAAGC-----	540
DB	714	GLInuARGbLeuLYbSERARGbLeuPROLYbTRPGLInTRPGLInLeuGLInuSERGLInuALbPGLA	733
QY	541	AGACGCTCTTAAAGAAAAGTTGCAAGAACAAACAAGC-----GAT	579
DB	734	ARGbInuSERLYbVALLeuLeuGLInuALbPROGLInuALbGLInuARGbALbGLInuALbPROGLInu	753
QY	580	TTAGAACAAAGAGACGCTCTTAAAGAAAAGTTGCA-----	615
DB	754	GINGLInGLInuLYbARGARGbARGbInuSERGLInuSERGLInuTRPGLInGLInGLInuARGbALbIS	773
QY	616	GAGCGACCAAGAGATTTAGAACAAAGGAAGCTGATACG---AAAAAAATTTAGAAAGA	672
DB	774	ARGbInuGLInGLInGLInuGLInuGLInuARGARGbPbPbInuTRPGLInuTRPGLInuALbGLInu	793
QY	673	AAAAAGGAACATGAGATATATTAGACAGAGATTTATATAGCGCT-----	717
DB	794	LYbSERGLInuARGbLY-----ARGbInuARGbLeuSERALbARGbPRObLeuLeuARGbGLIn	810
QY	718	-----TTAGAAATACACAGCTATAGAACTTCCATCGAATAAGAAAGCTGGATATTAT	768
DB	811	GINARGbInuARGbInuLeuARGbALbGLInuGLInuARGbInuGLInuARGbGLInuARGb-----Phe	828
QY	769	ATACCATCATCAATCTTCTTTACCTCAGGACCAACAGAGGGAATAGTAGAGATTTCAAAGCA	828
DB	829	LeuPROGLInGLInGLInuLYbSERGLInuARGbGLInuARGbInuARGbInuARGbGLInuARGbGLIn	848
QY	829	ATATCATATATAGAAAACAAATAG-----GAATCTTATACA	867
DB	849	LeuGLInuPbLeuGLInuGLInuGLInuGLInuGLInuGLInuGLInuGLInuGLInuGLInuGLIn	868
QY	869	ACAAAATGTTGAAGGA---CGAAGGATATATACATAAGACATCTTGAAGAAAAGAAAGAT	924
DB	869	GLInuGLInuGLInuPbLYbLeuGLInuGLInuPbGLInuGLInuARGARGARGbGLInuGLInuARGARGb	888
QY	925	GGTTCATATAAACAGACAAAGAAAGAGATTAATCTGCTGACATCAAAAATATATATCATTA	984
DB	889	GLInuLYbTRPARGbTRPGLInuLeuGLInuGLInuGLInuGLInuGLInuGLInuGLInuGLInu	906
QY	985	GAAGACGATTAATATTTCTGATGCTTATATGATTTCAATATAGTAAAGATAGAGAGATAAATA	1044
DB	907	-----TYRALbLYbPRObALbLeuGLInuGLInuGLInu	916
QY	1045	AGTGTGAATATGACGATTCATTAATAGATGAGAAAGAAAGATGATGAGACTTAGACGA	1104

QY 1357 CAGGTTAATAGAGAAAGAA 1377
Db 1141 LysPheArgGluGlnGln 1147
RESULT 13
ADM00404
ADM00404 standard; protein; 1439 AA.
XX
XX ADM00404;
XX
XX 07-APR-2005 (first entry)
XX
XX Epithelial hair follicle related antigen, SEQ ID 157.
XX
XX Endocrine-Gen.; Cardiovascular-Gen.; Venerary; Osteopathic; Cardiant;
XX Gene Therapy; antibody production; antibody therapy; antigen;
XX hair growth; arteriosclerosis; osteoporosis; bone injury; angina; burns;
XX wound healing.
XX
XX Mus musculus.
XX
XX US2005008602-A1.
XX
XX 13-JAN-2005.
XX
XX 07-JAN-2004; 2004US-00754079.
XX
XX 05-JUN-2000; 2000JP-00166903.
XX 04-JUN-2001; 2001WO-JP004691.
XX 13-NOV-2001; 2001JP-0034738.
XX 13-NOV-2001; 2001JP-00347340.
XX 05-DEC-2001; 2001JP-00371175.
XX 05-DEC-2001; 2001JP-00371366.
XX 23-MAY-2002; 2002JP-00155922.
XX 08-JAN-2003; 2003JP-00001891.
XX
XX (SUME) SUMITOMO ELECTRIC IND LTD.
XX
XX Hirai Y, Takebe K;
XX
XX MPI; 2005-080487/09.
XX
XX N-PsDB; ADM00406.
XX
XX New antibody specifically recognizing an antigen of 220kDa present in
XX epithelial new follicles, useful for evaluating and inducing hair growth
XX promoting activity and morphogenesis.
XX
XX Claim 5; SEQ ID NO 157; 87pp; English.
XX
XX The present invention relates to a novel antibody, or its fragment, which
XX specifically recognizes an antigen of about 220kDa (ADM00404) present in
XX new epithelial hair follicles. The antigen is specifically expressed
XX during the growth period of an imago or the developing period of a fetus.
XX Also claimed is a method for the evaluation of hair growth promoting
XX activity, comprising incubating skin tissue derived from a mammal in the
XX presence of a substance to be tested to promote hair growth, recovering
XX the skin tissue, reacting the skin tissue with the antibody, cited above,
XX and detecting the antibody, or its fragment, that reacts with the skin
XX tissue. The method and antibody of the invention are useful for
XX ameliorating and inducing hair growth promoting activity, treating and
XX inducing cardiovascular revascularization, regeneration and endotheelial
XX cell growth, such as in arteriosclerosis, Berger's disease, osteoporosis,
XX fractures and angina pectoris, and in treating burns or wounds.
XX
XX Sequence 1439 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1,4e-19 Length: 1439
XX Score: 374.00 Matches: 143
XX Percent Similarity: 48.15% Conservative: 130
XX Best Local Similarity: 25.22% Mismatches: 170
XX Query Match: 15.04% Indels: 124

DB: 9 Gaps: 25
US-09-837-344-41 (1-1482) x ADM00404 (1-1439)
QY 1 CAAGAACAAAGCGATCTTGAAACAAGAGAGCGTGTAAAGAAAG----- 48
Db 597 GlnGluGlnLysPheArgGlnLeuGlnGlnLysPheArgGlnGlnLysPheArgGln 616
QY 49 ---TTGCAAGAACAAAGCGATTTAGAACAAAGATGACTTGCTAAA-----GAA 96
Db 617 IleuGlnGlnGlnGlnPheGlnArgGlnIleuGlnGlnGlnLysPheArgGln 636
QY 97 AAGTTACAAAGACGCA----- 114
Db 637 ThrPheGlnGlnGlnGlnLysGlnLysGlnLysPheArgGlnGlnGlnLysGln 656
QY 115 ---AGCGATTTAGAACAAAGAGACTTCTTAAGAAAGTTGCAAGAACAAAGCGAT 171
Db 657 GlyLysPheLeuGlnGlnGlnLysPheArgGlnLysPheArgGlnGlnLysPheArg 676
QY 172 CTAGAACAAAGAGAGCGTGTAAAGAAAG-----TTGCAAGAACAAAGCGAT 222
Db 677 GlnGlnGlnGlnLysPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 696
QY 223 TTAGAACAAAGAGAGCGTGTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAA 282
Db 697 LeuArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 716
QY 283 GATAGACTTCTTAAGAAAGTTACAAAGCGAACAAAGCGATTTAGAACAAAGAGAG 339
Db 717 GlnGlnLysPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 735
QY 340 ---CGTGTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAG 390
Db 736 LeuGlnLysPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 755
QY 391 ---CGTGTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAG 441
Db 756 GlnGlnLysPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 775
QY 442 ---CTTGTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAG 492
Db 776 GlnGlnLysPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 795
QY 493 ---CGTGTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAG 543
Db 796 GlnGlnLysPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 815
QY 544 ---CGTGTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAG 597
Db 816 LeuGlnLysPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 835
QY 598 GCTTAAGAAAGTTGCAAGAGAGAACAAAGCGATTTAGAACAAAGCGATTTAGAA 657
Db 836 ---GlnGlnLysPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 853
QY 658 AAAAATTTAGAAAGAAAGAAAGCAATGAGATATATTTAGCAAGATTTATAGTCT 717
Db 854 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 777
QY 718 TTGAATATCCAGCTATAGACTTCCATCAGAAATGAAAGTGTATATATATCCACAT 777
Db 871 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 806
QY 778 CAATCTTTTATCTCAGAGACAAAGAGGAAATATGATGAT----- 819
Db 887 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 906
QY 819 ----- 819
Db 907 GlnLysPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 926
QY 820 TCCAGGAATATCTTATATAGAAAGAAAGCAATAGAGAG-----TCTATT 864

Dd		927	Hlervgtluueugluvalargtnglnleugluuagluurgleuaygarparglvbleu	946
Oy	ACAAACAATGTGGAAGCAGAAGCATATACATAAAGCACTTGGAAAAAAGAT	924		
Dd	947	Argargtgltugltuleuarlgarpar---	Arglvrbhenlvoglugtugluarg---	964
Oy	925	GATTCAATTAACCAAGAACAAAAGATMAATCTGCTGCATACAAATCAT-----		978
Dd	965	-----ArghlvgtlugluPhgltugltuvglneuluarglvneugluinProaparg		981
Oy	979	-----ACATTGAGACAGTAATATTTCTGATGTTAATGAT		1014
Dd	982	ArgPheargrgltugltuleuarlglnlcysvalglugltulaarglvleaargp		1001
Oy	1015	TTTTCAATTAAGTAGTAGAGATGAATAAGTGCTGAATATGACAGT-----		1062
Dd	1002	Serlysvlleargr---gluglncluleuargrgluarvgltugltugluarglvlea		1020
Oy	1063	-----TCATTATAGATGAAGAAAGAT		1086
Dd	1021	Aeparglvslleaargarpelngltuleuarlglnlgueneugltugltugltuleuar		1040
Oy	1087	GATBAAGACTTAGAC---GAATTTAAGCTTATGCGCAATATGACAAATTTCCAGATGAA		1143
Dd	1041	Arggtlngltuleuarparglvyphearggtugltugltuleuarpglngltuleugltu		1060
Oy	1144	GAANAACATA-----GGAATTTATTAAGACATGAGATGTTGATAGAG-----		1185
Dd	1061	Glurvgleuarparglvslleaargrgltugltugltuleuarlgarvlugltuleu		1080
Oy	1186	AAAAATGAAATTTAGATGATTTAGATGAAGAAATGAGAAATCATCAGAAATTAATCT		1245
Dd	1081	Argargrgltugltugltuleuararggtugltugltuleuararggtugltugltuleuar		1100
Oy	1246	GAAAGAAAAATTAAGAAAAAGAAAGAAATATGAAAAAAGATTAATATTTAAACA		1305
Dd	1101	Glutngltuleuarxngltuarxgtugltugltuleuarxgarparglvslleaarg		1120
Oy	1306	AATGAT-----AAAAGTTTGATGATGAGCATATTAAAAATTAATAATGATGAG		1356
Dd	1121	Aepgtngltuleuarxngltyleugltugltugltuleuarxgarvtugltugltuleuar		1140
Oy	1357	CAGGTTAATAAGAAAAAGAA		1377
Dd	1141	Lyerphearggtugltugltugltu		1147
RESULT 14				
ABM80133				
ID	ABM80133	standard; protein; 1584 AA.		
XX	ABM80133;			
XX	18-NOV-2004	(first entry)		
DE	Tumour-associated antigenic target (TAT) polypeptide PRO80641, SEQ:339.			
XX				
KW	Tumour-associated antigenic target; TAT; human; overexpression; cancer;			
KW	tumour; diagnosis; cell proliferative disorder; breast cancer;			
KW	colorectal cancer; lung cancer; ovarian cancer; liver cancer;			
KW	central nervous system cancer; bladder cancer; pancreatic cancer;			
KW	cervical cancer; melanoma; leukaemia; hybridisation probe;			
KW	chromosome identification; chromosome mapping; gene mapping;			
KX	gene therapy; cytostatic.			
XX				
OS	Homo sapiens.			
PN	WO2004030615-A2.			
PD	15-APR-2004.			
PP	29-SEP-2003; 2003WO-US028547.			
XX				

XX	02-OCT-2002: 2002US-041971P.
XX	(GENTH) GENENTECH INC.
PA	
XX	Wu TD, Zhang Z, Zhou Y;
P1	
XX	WPI: 2004-347921/32.
DR	N-PSDb; ACN37449.
XX	
PT	New tumor-associated antigenic target polypeptides and nucleic acids,
PT	useful in preparing a medicament for treating or detecting a
PT	proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT	prostate cancer or tumor.
PS	Claim 12; SEQ ID NO 339; 7273pp; English.
XX	
CC	The invention relates to human tumour-associated antigenic target (TAT)
CC	polypeptides, and their related nucleic acids. The TAT polypeptides are
CC	overexpressed in cancer tissues compared to normal tissues, and may thus
CC	serve as effective targets for the diagnosis and treatment of cancer in
CC	mammals. The invention also relates to nucleic acid and polypeptide
CC	sequences at least 80% identical to the TAT nucleic acid and
CC	polypeptides; expression vectors and host cells comprising a TAT nucleic
CC	acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC	molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC	TAT polypeptide; and methods and compositions for the treatment or
CC	diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC	antibodies, antagonists, binding molecules and compositions are useful
CC	for diagnosing or treating a cell proliferative disorder associated with
CC	increased TAT expression, particularly cancers such as breast cancer,
CC	colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC	cancer, pancreatic cancer, cervical cancer, cancers of the central
CC	nervous system, melanoma and leukemia. TAT nucleic acids may further be
CC	used as hybridisation probes, in chromosome and gene mapping, in
CC	chromosome identification and in gene therapy. The present sequence
XX	represents a TAT polypeptide of the invention
XX	
SQ	Sequence 1584 AA:
	Alignment Scores:
	Pred. No.: 1.68e-19 Length: 1584
	Score: 373.00 Matches: 133
	Percent Similarity: 46.53% Conservative: 115
	Best Local Similarity: 24.95% Mismatches: 185
	Query Match: 15.00% Indels: 100
	DB: 8 Gaps: 21
	US-09-837-344-41 (1-1482) x AEW80133 (1-1584)
OY	1 CAAGAACAAACAAGCGCATCTAGAACACAGAGAGCGTGCTTAAGAAAAGTTGCAGAACA 60
Db	278 ArgGIuGIuGIuInIeuArGIuArGIuGIuGIuArGIuArGIuArGIuArGIuIn 29
OY	61 CAAAGCGATTTAGAACAAGATAGACTGTCTTAAGAAAAGTTTCAMAGACGACAGCAAGCAT 120
Db	298 GIu-----GIuGIuGIuArGIuArGIuGIuInArGIuArGIuArGIuGIuInIeuArg 31
OY	121 TTAGAACAAAGAGACTTGCTTAAGAAAAGTTG---CAAGAACAAACAAGCATTAGAA 17
Db	315 ArgGIuGIuGIuGIuArGIuArGIuGIuInIeuArGIuArGIuGIuGIuArGIuIn 33
OY	178 CAAAGAGAGCGTGCTTAAGAAAAGTTGCAGAACAACAAGC----- 21
Db	335 GIuGIuArGIuArGIuGIuInArGIuArGIuGIuGIuInIeuArGIuArGIuArGIuIn 35
OY	220 GATTTAGAACAAAGAGACCGTGCTTAAGAAAAGTTGCAGAACAACAAGCATTAGAA 27
Db	355 LysArGIuGIuGIuArGIuArGIuGIuInArGIuArGIuGIuGIuInIeuArGIuIn 37
OY	280 CAAGATGACTTGCTTAAGAAAAGTTACAA-----GAGCAGCAAAGCATTTAGAACAA 33
Db	375 GIu--ArGIuLysArGIuGIuGIuGIuArGIuArGIuGIuGIuGIuArGIuGIuIn 39

```

QY 334 GAGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTA-----GAA 381
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 334 ArgLeuIysArgGluGluProGluIuArgGlnIuLeuIysSerGlu 413
QY 382 CAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAAGACAGAGA 441
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 GlnGluIuArgArgGlnIuGlnIuLeuArgArgGlnGlnIuArgArgGlnIuArg 433
QY 442 CTTCGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAAGACAA 486
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 434 LeuIysArgGluGluGlnIuGlnIuArgGlnIuArgLeuIysArgGlnIuIuGln 453
QY 487 GAGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAAGACAGAG 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 454 GluArgArgGluGlnIuGlnIuLeuIuArgGlnIuGlnIuArgGlnIuArg 473
QY 526 -----GATTTAGAACAGAG----- 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 SerArgIleProIysTrpGlnTrpGlnLeuGlnSerGlnIuIuAspAlaArgGlnSerIys 493
QY 541 -----AGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAAGACAGAG 591
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 494 ValIyrSerArgProArgIysGlnIuGlnIuArgArgArgGlnIuGlnIuIuIys 513
QY 592 AGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAAGACAGAGCA 624
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 514 ArgArgArgArgArgGlnSerGlnIuGlnIuArgGlnIuGlnIuArgAlaIuIuArgGln 533
QY 625 AGACGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAAGACAGAGCA 681
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 534 GlnGlnIuGlnIuGlnIuArgArgArgPheThrTrpGlnTrpGlnIuIuArgGlnIuIys 553
QY 682 CATGAGATATATTAGCAGACGATTATATGTCGT----- 717
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 554 ArgGly-----ArgGlnArgLeuSerAlaArgProIuLeuArgGlnIuArgGlu 570
QY 718 TTAGAAATACAGCTATAGAACTTCATCAGAAAAGCGATATTATATACACAT 777
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 571 ArgGlnLeuArgAlaGlnIuArgGlnIuArgGlnIuArg-----PheLeuProGlu 588
QY 778 CAATCTCTCTTACCTCAGCAACAGAGGAAATAGTAGAATTCAGAGAAATATCTATA 837
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 589 GlnGlnIuGlnIuGlnIuArgArgArgGlnArgGlnIuArgGlnIuIysLeuGlnPhe 608
QY 838 ATAGAAAAAACAATAGAGATCTATTACAAACAAAGTTGAAAGACGAAAGGATATACAT 897
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 609 LeuGlnIuGlnIuGlnIuLeuGlnArgArgGlnArgAlaGlnIuLeuGlnIuGlnIu 628
QY 898 AAAGGACATCTTGAAGAAAAGATGGTTCAATTAACCAAGCAACAAAGAAAGATAAA 957
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 629 AspGly-----LeuGlnIuAspGlnIuArgArgArgSerGlnIuGlnIuArgArgAspGln 647
QY 958 TCTGCTGACATACAA-----AATCATTCATTAGACACAGTA 993
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 648 LysTrpArgTrpGlnIuGlnIuGlnIuArgIysArgArgArgHisThrLeu----- 664
QY 994 AATATTTCGATCTTAATGATTTTCAATAAGTAAGTAAGAGATTAATAGTGTCTGAA 1053
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 665 -----TyrAlaIysProAlaLeuGlnIuGlnIuLeuArgIysGln 677
QY 1054 TATGACGATTCATTATAGTAGAAGAAAGATGATGAACTTAAGCAATTTAAGCCT 1113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 678 -----GlnGlnIuLeuGlnIuGlnIuGlnIuGlnIuArgGlnIu----- 692
QY 1114 ATTGTCGATATAGCAATTTCCAAAGATGAGAAAACATAGGAATTTATAAGAACTAGAA 1173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 693 -----ArgGlnIuIysArgArgArgGlnIuGlnIuArgGlnIuIysArgGlnIuGln 709
QY 1174 GATTTGATAGAGAAAATGAAAATTTAGATGATTTAGATGAGAGAAATAGAAAATATCATCA 1233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 710 GlnLeuGlnIuGlnIuGlnIuGlnIu-----LeuArgGlnIuArgGlnIuIysArgArg 727
QY 1234 GAAGAATTATCTGAGAAAAAATAAAGAAAAGAAAGAAATATATGAAAAAACAAGATTAAT 1293

```

```

Db 728 ArgGlnIuArgGlnIuArgGlnIuIysAspIysLeuGlnIuIysGlnIuGln 747
QY 1294 AATTTT---AACAACAAATGATTAAGTTGATGATGACATATTAAAAAATATATAAT 1350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 748 LeuLeuGlnIuGlnIuProGlnIuIysArgArgArgGlnIuArgGlnIuIysArgGln 767
QY 1351 GATAAGCAGTTTAATTAAGAAAAGAAAAGAAAATTCATTAATA 1389
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 768 GlnGlnIuLeuGlnIuGlnIuGlnIuGlnIuLeuLeuArg 780

RESULT 15
AAW03626
ID AAW03626 standard; peptide, 412 AA.
AC AAW03626;
XX 25-MAR-2003 (revised)
DT 17-DEC-1996 (first entry)
XX
DE Human thyrotropin GPR N-terminal sequence.
KM G-protein coupled receptor; ligand binding assay; transmembrane domain;
KM schizoprenia; dopamine; cAMP; adenosine; thrombin; adrenergic; opsin;
KM muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin;
KM odorant; cytomegalovirus; serotonergic.
OS Homo sapiens.
XX
XX US508384-A.
XX
XX 16-APR-1996.
XX
XX 09-SEP-1993; 93US-00118270.
XX
XX 10-SEP-1992; 92US-00943236.
XX
XX (UNYV ) UNIV NEW YORK STATE.
XX
XX Schuster DI, Murphy RB;
XX
XX WPI, 1996-208785/21.
XX
XX
XX New dopamine receptor peptide - useful as antipsychotic agent, e.g. for
XX treating schizophrenia.
XX
XX Disclosure; Fig 8B(2); 184pp; English.
XX
XX Proteins AAW02657-W02720 represent a range of G-protein coupled receptor
XX (GPR) proteins selected from cAMP, adenosine, muscarinic acetylcholine,
XX adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin,
XX odorant, cytomegaloviral and other GPR proteins. The peptides AAW03578-
XX W03581 represent the N-terminal fragments of the above proteins. The
XX receptor proteins were used to design polypeptides, pref. based on the
XX transmembrane domains, for use in G-protein coupled receptor ligand
XX binding assays. The polypeptide fragments retain biological activity such
XX as binding a GPR ligand or modulating GPR ligand binding to a GPR (see
XX AAW02747-W02999 for examples of polypeptide fragments). The polypeptide
XX fragments can be used in compositions for treating subjects suffering
XX from a pathology related to a GPR abnormality e.g. a psychotic disorder
XX such as schizophrenia. (Updated on 25-MAR-2003 to correct PF field.)
XX
XX
SQ Sequence 412 AA;

Alignment Scores:
Pred. No.: 1,04E-18 Length: 412
Score: 361.50 Matches: 80
Percent Similarity: 51.73% Conservative: 129
Best Local Similarity: 19.80% Mismatches: 140
Query Match: 14.54% Indels: 55
DB: Gaps: 5

US-09-837-344-41 (1-1482) x AAW03626 (1-412)

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 7, 2005, 06:33:52 ; Search time 31.5 Seconds

(without alignments)
7779.388 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 2486
Sequence: 1 CAAACAACACAAAGCCATCT.....AATATTATGAACATATA 1482

Scoring table:
BLSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=x1h
-O=cgnt2_1/USPTO.apool.b/US09837344/runat.05122005.133407.29484/app.query.fasta_1.1671
-DB=Issued_Patents_AA -OPMT=faetan -SUFFIX=rai -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09837344@CGN 1.1.34@runat.05122005.133407.29484 -NCPU=6 -ICPU=3
-NO MMAP -IARGOUTDIRY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

Database: Issued Patents AA:

1: /cgnt2_6/ptodata/1/1aa/5/COMB.pep:*
2: /cgnt2_6/ptodata/1/1aa/6/COMB.pep:*
3: /cgnt2_6/ptodata/1/1aa/H/COMB.pep:*
4: /cgnt2_6/ptodata/1/1aa/PTUS/COMB.pep:*
5: /cgnt2_6/ptodata/1/1aa/RE/COMB.pep:*
6: /cgnt2_6/ptodata/1/1aa/backfill1est.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1549	62.3	316 2	US-08-098-327E-31 Sequence 31, Appl
2	1549	62.3	316 2	US-08-462-625-31 Sequence 31, Appl
3	603	24.3	117 2	US-08-098-337E-20 Sequence 20, Appl
4	603	24.3	117 2	US-08-462-625-20 Sequence 20, Appl
5	554	22.3	107 2	US-08-098-337E-19 Sequence 19, Appl
6	554	22.3	107 2	US-08-462-625-19 Sequence 19, Appl
7	376	15.1	1898 1	US-08-056-200-94 Sequence 94, Appl
8	376	15.1	1898 1	US-08-800-644-94 Sequence 94, Appl
9	376	15.1	1898 2	US-09-538-092-1280 Sequence 1280, Ap
10	359	14.4	1162 1	US-08-728-333A-2 Sequence 2, Appl
11	359	14.4	1162 1	US-09-298-568-2 Sequence 2, Appl
12	359	14.4	1162 2	US-09-410-399-2 Sequence 2, Appl

13	359	14.4	1162 2	US-09-894-273-2 Sequence 2, Appl
14	336	13.5	608 2	US-09-270-767-32937 Sequence 32937, A
15	336	13.5	608 2	US-09-270-767-48154 Sequence 48154, A
16	323.5	13.0	611 2	US-09-216-393B-81 Sequence 81, Appl
17	319	12.8	1972 2	US-08-875-435B-3 Sequence 3, Appl
18	317.5	12.8	1937 2	US-09-538-092-918 Sequence 918, App
19	316	12.7	1180 2	US-09-543-681A-6436 Sequence 6436, Ap
20	315.5	12.7	1564 2	US-10-144-198-2 Sequence 2, Appl
21	315.5	12.7	1564 2	US-10-144-198-4 Sequence 4, Appl
22	315	12.7	1960 2	US-09-538-092-1077 Sequence 10872, A
23	315	12.7	1960 2	US-09-849-016-10872 Sequence 1078, A
24	314	12.6	1976 2	US-09-538-092-1078 Sequence 4, Appl
25	309	12.4	1972 2	US-08-875-435B-4 Sequence 8155, Ap
26	306.5	12.3	1942 2	US-09-849-016-8115 Sequence 915, App
27	305	12.3	1939 2	US-09-538-092-915 Sequence 11104, A
28	305	12.3	1939 2	US-09-849-016-11104 Sequence 6925, Ap
29	304.5	12.2	1939 2	US-09-949-016-6925 Sequence 18798, A
30	303.5	12.2	1939 2	US-09-248-796A-18798 Sequence 917, App
31	303	12.2	1939 2	US-09-310-187A-1 Sequence 917, App
32	303	12.2	1939 2	US-09-538-092-917 Sequence 8134, Ap
33	302	12.1	1959 2	US-09-817-254-91 Sequence 91, Appl
34	301	12.1	1857 2	US-09-538-092-1084 Sequence 1084, Ap
35	301	12.1	1972 2	US-09-538-092-1084 Sequence 7112, Ap
36	301	12.1	1984 2	US-09-949-016-7112 Sequence 7113, Ap
37	301	12.1	1984 2	US-09-949-016-7113 Sequence 916, App
38	301	12.1	1984 2	US-09-949-016-7113 Sequence 11, Appl
39	299.5	12.0	1886 2	US-09-538-092-916 Sequence 10929, A
40	299.5	12.0	1935 2	US-09-914-259-11 Sequence 80, Appl
41	299	12.0	1847 2	US-09-949-016-10929 Sequence 79, Appl
42	298.5	12.0	1944 2	US-10-164-595-80 Sequence 54, Appl
43	298	12.0	784 2	US-10-164-595-79 Sequence 54, Appl
44	298	12.0	843 2	US-10-164-595-54 Sequence 54, Appl
45	298	12.0		

ALIGNMENTS

RESULT 1
US-08-098-327E-31
; Sequence 31, Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUESIN-MARCHAND, Claudine
; APPLICANT: DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098.327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620

Db 221 Lybysgclunhsglyaspielenuaiaaapueutyrglyargueugliuleprola 240
Gy 733 ATGAACCTTCATCAGAAAATGAACGTGGATTTATATACACATCAATCTTCTTAAGT 792

DB: 2 Ga

Pred. No.:	1.09e-117	length:	316
Score:	154.9	Matches:	314
Percent Similarity:	99.37%	Conservative:	0
Best Local Similarity:	99.37%	Mismatches:	2
Query Match:	62.31%	Indels:	0
DB:	2	Gaps:	0


```

QY 910 GAAGAAAGAAAGATGTTTCATATAAACCGAACAAGAAAGAAAGTAAATCT 960
Db 101 GUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 117

RESULT 4
US-08-462-625-20
; Sequence 20, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; NUMBER OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,625
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/098,327
; FILING DATE: 24-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
;
US-08-462-625-20

Alignment Scores:
Pred. No.: 4,03e-41 Length: 117
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.26% Indels: 0
DB: 2 Gaps: 0

US-09-837-344-41 (1-1482) x US-08-462-625-20 (1-117)
QY 610 TTGCAAGACGACGACAAAGATTAGAACAAAGAAAGCTGATACGAAAAAATTAGAA 669
Db 1 LGGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 20
QY 670 AGAAAAAGAAACATGAGATATATTTCGACAGAGATTATATGTGCTTTGAAATACCA 729
Db 21 ATGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 40

```

OY	730	GCCTATAACTTCCATCAGAAAAATGACCGTGATATATATACACATCAATCTTCTTA	789
Db	41	AlAllegluueuProSerGluAenGluArgIlyTyTyrlleProHisGlnSerIeu	60
OY	790	CCTCAGACACACAGAGGGAATAGTAGAGATTCCAGGAAATATCTTATATAGAAAAACA	849
Db	61	ProGlnAspAsnArgGlyAenSerArgAspSerIeuIleSerIleIleGluIyThr	80
OY	850	AATGAGAAATCTATTACACAAATGTTGAAGACGAAAGGATATACATAAGGACATCTT	909
Db	81	AsnArgGluSerIleThrIleAsnValGluGlyAArgArgAspIleHisIlysgIyHisIeu	100
OY	910	GAACAAAAAGAAAGATGGTTCATTAACCAACGACAAAAAGAGATAATCT	960
Db	101	GluGluIyIyIyAspGlySerIleIyIyProGluGlnIyIyGluIyAspIySer	117
RESULT 5			
US-08-098-327E-19			
Sequence 19, Application US/08098327E			
Patent No. 6270771			
GENERAL INFORMATION:			
APPLICANT: GUERIN-MARCHAND, Claudine			
APPLICANT: DRULHE, Pierre			
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE			
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE			
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES			
NUMBER OF SEQUENCES: 46			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Burns, Doane, Swecker & Mathis			
STREET: P.O. Box 1404			
CITY: Alexandria			
STATE: Virginia			
COUNTRY: United States			
ZIP: 22313-1404			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/098.327E			
FILING DATE: 24-NOV-1993			
CLASSIFICATION: 424			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: FR 91 01286			
FILING DATE: 05-FEB-1991			
ATTORNEY/AGENT INFORMATION:			
NAME: McGowan, Malcolm K.			
REGISTRATION NUMBER: 39,300			
REFERENCE/DOCKET NUMBER: 010830-045			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (703) 836-6620			
TELEFAX: (703) 836-2021			
INFORMATION FOR SEQ ID NO: 19:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 107 amino acids			
TYPE: amino acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: peptide			
PUBLICATION INFORMATION:			
DOCUMENT NUMBER: WO 92/13884			
PUBLICATION DATE: 20-AUG-1992			
US-08-098-327E-19			
Alignment Scores:			
Pred. No.: 3.71e-37			
Score: 554.00			
Percent Similarity: 100.00%			
Best Local Similarity: 100.00%			
Query Match: 22.28%			
DB: 2			
Length: 107			
Matches: 107			
Conservative: 0			
Mismatches: 0			
Indels: 0			
Gaps: 0			

US-09-837-344-41 (1-1482) x US-08-098-327E-19 (1-107)

QY 640 AGGAGGCTGATACGAAAAAATTTAGAAAAAGAACATGAGATATATTACCA 699

Db 1 ArglyeAlaaphrlyslslysaenleuGluArglyslslyslleuAla 20

QY 700 GAGATTTATATGCTGTTTGAATTCAGCTTTAGAACTTCCATCAGAAAAATGAACGT 759

Db 21 GluAspLeuTyrglyArgleuGluileProAlalleuLeuProserGluAenGluArg 40

QY 760 GATATATATACACATCAATCTTTTACCTCAGACACAGAGGAATAGTAGAGAT 819

Db 41 GlyTyrtlyrileProhlsGlnSerSerleuProGlnAspAsnArglyAenSerArgAsp 60

QY 820 TCCAGGAAATATCTATATAGAAAAACAATAGAAATCTATTAACAATGTTGAA 879

Db 61 SerlyseGluileserilelleGluTyethrAsnArgGluSerilethrThrAenValGlu 80

QY 880 GAGCAAGGATATACATAAGACATCTTGAGAAAAAGAAATGCTTCAATTAACCA 939

Db 81 GlyArgAspIleHislyslslyslsleuGluGluTyAspGlySerileTyPro 100

QY 940 GAACAAAAAGAGATAATCT 960

Db 101 GluGlnlyslslyAspLysSer 107

RESULT 6

US-08-462-625-19

Sequence 19 Application US/08462625

Patent No. 6319502

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

APPLICANT: DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

NUMBER OF INVENTION: OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,625

FILING DATE: 05-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/098,327

FILING DATE: 24-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 91 01286

FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

US-08-462-625-19

Alignment Scores:

Prod. No.: 3,71e-37

Score: 554.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 22.28%

DB: 2

Gaps: 0

Length: 107

Matches: 107

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-837-344-41 (1-1482) x US-08-462-625-19 (1-107)

QY 640 AGGAGGCTGATACGAAAAAATTTAGAAAAAGAACATGAGATATATTACCA 699

Db 1 ArglyeAlaaphrlyslslysaenleuGluArglyslslyslleuAla 20

QY 700 GAGATTTATATGCTGTTTGAATTCAGCTTTAGAACTTCCATCAGAAAAATGAACGT 759

Db 21 GluAspLeuTyrglyArgleuGluileProAlalleuLeuProserGluAenGluArg 40

QY 760 GATATATATACACATCAATCTTTTACCTCAGACACAGAGGAATAGTAGAGAT 819

Db 41 GlyTyrtlyrileProhlsGlnSerSerleuProGlnAspAsnArglyAenSerArgAsp 60

QY 820 TCCAGGAAATATCTATATAGAAAAACAATAGAAATCTATTAACAATGTTGAA 879

Db 61 SerlyseGluileserilelleGluTyethrAsnArgGluSerilethrThrAenValGlu 80

QY 880 GAGCAAGGATATACATAAGACATCTTGAGAAAAAGAAATGCTTCAATTAACCA 939

Db 81 GlyArgAspIleHislyslslyslsleuGluGluTyAspGlySerileTyPro 100

QY 940 GAACAAAAAGAGATAATCT 960

Db 101 GluGlnlyslslyAspLysSer 107

RESULT 7

US-08-056-200-94

Sequence 94 Application US/08056200

Patent No. 5616500

GENERAL INFORMATION:

APPLICANT: Steinert, Peter M.

APPLICANT: Lee, Seung-Chul

APPLICANT: Kim, In-Gyu

APPLICANT: Chung, Soo-Il

TITLE OF INVENTION: Trichoyallin and Transglutaminase-3 and

TITLE OF INVENTION: Methods of Using Same

NUMBER OF SEQUENCES: 117

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 620 Newport Center Drive, Sixteenth Floor

CITY: Newport Beach

STATE: CA

COUNTRY: U.S.A.

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/056,200

FILING DATE: 30-APR-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fedrick, Michael P.

REGISTRATION NUMBER: 36,799

REFERENCE/DOCKET NUMBER: NIH054.001A


```
Db      931 GIUthrValGIUGluPrOtle 937

RESULT 11
US-09-298-568-2
Sequence 2, Application US/09298568
Parent No. 6322792
GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Ballesteras, Mary E.
APPLICANT: Kaye, Kenneth M.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
FILE REFERENCE: 16412-1000IR
CURRENT APPLICATION NUMBER: US/09/298,568
CURRENT FILING DATE: 1999-04-21
EARLIER APPLICATION NUMBER: US 60/109,422
EARLIER FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1162
TYPE: PRT
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-2

Alignment Scores:
Pred. No.:          3,2e-21          Length:          1162
Score:              359.00           Matches:           94
Percent Similarity: 49.44%           Conservative:     127
Best Local Similarity: 21.03%         Mismatches:       110
Query Match:        14.44%           Indels:            16
DB:                  2                Gaps:             12

US-09-837-344-41 (1-1482) x US-09-298-568-2 (1-1162)
QY      4 GAACACAAAGCGATTGTAACAAGAAGACCGCTCAAAAAGAAAAAGTTGCAGAACAACA 63
        |||||
Db      604 GIUGInGIInIasp---GIUGInGIInIaspGIUGInGIInIaspGIUGInGIInI 622
               :|||||
QY      64 AGCGATTTAGAACAAAGTGAAGACTTGCTAAAGAAAAGTTACAA-----GAGCAGCAAAGC 117
        |||||
Db      623 AepGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInI 642
               :|||||
QY      118 GATTTAGAACAAGAAGACCTGCTAAAGAAAAGTTGCAGAACAACAAGCGACTAGAA 177
        |||||
Db      643 AepGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInI 662
               :|||||
QY      178 CAAGAGACGCGCTTAAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAGAAGA 237
        |||||
Db      663 AepGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInI 682
               :|||||
QY      238 CGTCGTAAAGAAAAGTTGCAGAACAACAAGCGATTTAGAACAGATGACTTGCTAAA 297
        |||||
Db      683 AepGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInI 702
               :|||||
QY      298 GAAAAGTTACAAGACGACCAAGCGATTTAGAACAGAAGACGTGCTAAAGAAAAGTTG 357
        |||||
Db      703 GIaepGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInI 722
               :|||||
QY      358 CAAAGAACAAAGCGATTTAGAACAGAAGACGCGCTTAAAGAAAAGTTGCCAAGAAC 417
        |||||
Db      723 GIInGIInGIInIaspGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInI 742
               :|||||
QY      418 CAAAGCATTTAGAACAGAAGACACTTGTCAAGAAAAGTTGCAGAACAACAAGCGAT 477
        |||||
Db      743 GIInIaepGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInIaspGIUGIn 758
               :|||||
QY      478 TTAGAACAAAGACGCTGCTAAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAA 537
        |||||
Db      759 GIInGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInIaspGIUGInGIInI 778
               :|||||
QY      538 GAGAGACGTGCTAAAGAAAAGTTGCCAAGAACAAAGCGATTTAGAACAGAAGACGT 597
        |||||
```

[illegible]

Db 643 AepglunglInglInaepglunglInglInaepglunglInglInaepglunglInglIn 662
QY 178 CAAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGGCGATTGTGAACAAGAGAGA 237
Db 663 AepglunglInglInaepglunglInglInaepglunglInglInaepglunglInglIn 682
QY 238 CGTGTAAAGAAAGTTGCAAGAACAAAGGCGATTGTGAACAAGATAGCTTCTAA 297
Db 683 AepglunglInglInaepglunglInglInaepglunglInglInaepglunglInglIn 702
QY 298 GAAAAGTTAACAAGACGCAAAAGCGATTGTGAACAAGAGAGCGCTTAAAGAAAGTTG 357
Db 703 GlnaepglunglInglInaepglunglInglInaepglunglInglInaepglunglIn 722
QY 358 CAAGAACCAAGGCGATTGTGAACAAGAGAGCGCTTAAAGAAAGTTGCAAGACAA 417
Db 723 GlnGlnGlnGlnaepglunglInglInaepglunglInglInaepglunglInglIn 742
QY 418 CAAGGCGATTGTGAACAAGAGAGCTTGTAAAGAAAGTTGCAAGAACAAAGCGCAT 477
Db 743 GlnGlnaepglunglInglIn-----GlnaepglunglInglInglInglIn 758
QY 478 TTGAACAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGTGAACAA 537
Db 759 GlnGlnGlnGlnInglInglInglInglInglInglInglInglInglInglInaep 778
QY 538 GAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGTGAACAAGAGAGACT 597
Db 779 Gln-----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln----- 793
QY 598 GCTAAAGAAAGTTGCAAGAGCAGCAAAAGAGATTGTGAACAAGAAAGCGCTATACGAA 657
Db 794 ---GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln-----Gln 808
QY 658 AAAAATTGTGAAGAAAGAAAGCAATGAGATATATTAGCAGAGATTATATGCTGCT 717
Db 809 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 817
QY 718 TTGAAGAAATCCAGCTATAGAACTTCATCAGAAATGAACGCGGATTTATATACCACT 777
Db 817 ----- 817
QY 778 CAATCTTCTTAACTCAGAGCAACAGAGGGAATAGAGAGATTTCAAGGAAATATCTATA 837
Db 817 ----- 817
QY 838 ATAGAAAAAACAATAGAGATCTATTACAACAATGTTGAAGAGCAAGGATATATCAT 897
Db 818 LæuglInglInglInglIn----- 824
QY 898 AAAGGCACTTGTGAAGAAAGAAAGTGTCAATTAACCAAGAACAAAGAAAGATAA 957
Db 825 -----LæuglInglInglInglInglInglInglInglInglInglIn 837
QY 958 TCTGCTGACATACAAATCATATCATTAGAGACAGTAATATTTCTATGTTAATGATTT 1017
Db 838 -----GlnLæuglInglIn 842
QY 1018 CAATATAGTAAGTAGAGATGAATAGCTGAATATGACGATTCATTAATAGATGA 1077
Db 843 GlnValGlnGlnGlnGlnGlnValGlnGln-----GlnGln 856
QY 1078 GAAGAAAGTAGTAAGACTTATAGCAATTTAAGCTTATGCAATATGCAATTTCCAA 1137
Db 857 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 876
QY 1138 GATGAAGAAACATAGGAATTTATAAGACTGAAGATTTGATAGCAAAATGAA----- 1194
Db 877 GlnGlnGln-----GlnLæuglInglInValGlnGlnGlnGlnGlnGlnGln 890
QY 1195 AATTATGATGATTTAGATGAAGATGAAGAAATTCATCAGAGAAATTTATCGAAGAAAA 1254
Db 891 GlnLæuglInglInValGlnGlnGlnGlnGlnGlnGlnLæuglInglInValGlnGlnGlnGln 910

QY 1255 ATAAAAAAGAAAGAAATATGAAAAAACAAGATTAATTTTAAACCAATGATATA 1314
Db 911 GlnGlnGlnLæuglInglInglInValGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 930
QY 1315 AGTTTGTATGATGACCATATT 1335
Db 931 GlnThrValGlnGlnProfile 937
RESULT 14
US-09-270-767-32937
; Sequence 32937, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32937
; LENGTH: 608
; TYPE: PRN
; ORGANISM: Drosophila melanogaster
US-09-270-767-32937
Alignment Scores:
Pred. No.: 2,146-19 Length: 608
Score: 336.00 Matches: 70
Percent Similarity: 64.94% Conservative: 80
Best Local Similarity: 30.30% Mismatches: 73
Query Match: 13.52% Indels: 8
DB: 2 Gaps: 4
US-09-837-344-41 (1-1482) x US-09-270-767-32937 (1-608)
QY 1 CAAGAACAAACAAAGCGATCTAGAACACAGAGACGTGCTAAAGAAAGTTGCAAGAACAA 60
Db 374 LysGlnGlnGlnGlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 393
QY 61 CAAGCGATTGTGAACAAGATAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 120
Db 394 GlnArgArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 413
QY 121 TTGAACAAGAGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
Db 414 LysGlnLysArgArgArgArgArgArgArgArgArgArgArgArgArgArgArgArgArg 433
QY 181 GAGAGCGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGT 240
Db 434 GlnArgLysGlnGlnGlnGlnArgArgArgArgArgArgArgArgArgArgArgArgArg 452
QY 241 GCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCGTCTTAAAGAA 300
Db 453 AlaGlnGlnArgArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 472
QY 301 AAGTTAACAAGAGAGCAAGCGATTGTGAACAAGAGAGCGCTTAAAGAAAGTTGCA 360
Db 473 LysArgArgArgArgGln-----GlnGlnArgArgLysGlnGlnGlnGlnGlnGlnGln 489
QY 361 GAACAAACAAGCGATTGTGAACAAGAGAGCGTCTTAAAGAAAGTTGCAAGAACAA 417
Db 490 GlnGlnGlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 509
QY 418 CAAGCGATTGTGAACAAGAGAGCTTGTCTAAAGAAAGTTGCAAGAACAAAGCGAT 477
Db 510 GlnLysArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 529
QY 478 TTGAACAAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAA 537
Db 530 LysGlnGlnGlnLysArgArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 549

```
QY 538 GAGAGACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAGAGACGT 597
|||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 550 GluArgArgGluLysGluArgArgGluArgGluArgGluArgGluArgGlu 569
QY 598 GCTAAAGAAAAGTTGCAAGACGACAAAGCGATTAGAACAA-----AGGAGGCT 648
::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|
Db 570 GluGluGluArgArgGluGluGluArgArgGluGluArgGluGluArgGlu 589
QY 649 GATACGAAAAAAATTTAGAAAAGAAAAGGAA 681
::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|
Db 590 GluArgArgArgGluGluGluArgLysArgGlu 600

RESULT 15
US-09-270-767-48154
; Sequence 48154, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 1999-03-17
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 48154
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48154

Alignment Scores:
Pred. NO.: 2,14e-19 Length: 608
Score: 336.00 Matches: 70
Percent Similarity: 64.94% Conservative: 80
Best Local Similarity: 30.30% Mismatches: 73
Query Match: 13.52% Indels: 8
DB: 2 Gaps: 4

US-09-837-344-41 (1-1482) x US-09-270-767-48154 (1-608)
QY 1 CAAGAACACAAAGCGATCTAGAACAGAGACGTCTAAAGAAAAGTTGCAAGAACAA 60
::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|
Db 374 LysGluGluGluLysGluArgGluGluArgGluArgGluArgGluArgGlu 393
QY 61 CAAAGCGATTAGAACAGATGACTTCTAAAGAAAAGTTAGAACAGACGACAAAGCGAT 120
::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|
Db 394 GluArgArgGluGluGluGluArgGluArgGluArgGluArgGluArgGlu 413
QY 121 TTGAAACAAAGAGACCTTCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAA 180
|||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 414 LysGluLysArgArgArgGluLysArgGluLysArgGluLysArgGluLys 433
QY 181 GAGAGACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGT 240
|||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 434 GluArgLysGluGluGluArgGluArgGluLysArgGluLysArgGluLys 452
QY 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGATGACTTCTAAAGAAA 300
|||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 453 AlaGluGluArgArgGluLysGluGluArgGluArgGluLysArgGluArg 472
QY 301 AAGTTACAAAGCAGCAAGCGATTTAGAACAGAGACGTCTAAAGAAAAGTTGCA 360
|||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 473 LysArgArgArgGluGlu-----GluGluArgArgGluGluArgGlu 489
QY 361 GAACAACAAAGCGATTTAGAACAGAGACGTCTCTCTCTCTCTCTCTCTCTCTCTCT 417
|||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 490 GluGluGluLysArgGluGluGluGluArgGluArgGluArgGluArgGlu 509
QY 418 CAAAGCGATTTAGAACAGAGACGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 477
::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|
Db 510 GluLysArgGluGluGluArgGluLysArgGluArgGluArgGluArgGlu 529
QY 478 TTAGAACAGAGACGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 537
```

```
|||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 530 LysGluGluGluArgGluLysGluArgArgGluGluGluLysArgGlu 549
QY 538 GAGAGACGTCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGT 597
|||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 550 GluArgArgGluLysGluArgArgGluGluArgArgGluGluArgArgGlu 569
QY 598 GCTAAAGAAAAGTTGCAAGACGACAAAGCGATTAGAACAA-----AGGAGGCT 648
::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|
Db 570 GluGluGluArgArgArgGluGluGluArgArgArgGluGluArgArgGlu 589
QY 649 GATACGAAAAAAATTTAGAAAAGAAAAGGAA 681
::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|  ::|::|
Db 590 GluArgArgArgGluGluGluArgLysArgGlu 600
```

Search completed: December 7, 2005, 06:57:51
Job time : 55.5 secs

THIS PAGE BLANK (USPTO)

QY 88 -----GCTTAAAGAAAGTTACAGAGCAGCAAGGATTTAGAAACA 129
Db LeuGInGInuArGArGAlaLySGInuLySGInGInGInSeArSpLeuGIn 1425
QY 130 GAGACACTTGCTTAAAGAAAGTTGCAAGAACAAACGATCTAGAACAGAGAGAGCT 189
Db 1426 GluArGArGAlaLySGInuLySGInGInGInSeArSpLeuGInGInuArGArG 1445
QY 190 GCTTAAAGAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGAGCTGCTTAAAGAA 249
Db 1446 AlaLySGInuLySGInGInGInSeArSpLeuGInGInuArGArGAlaLySGIn 1465
QY 250 AAGTTGCAAGAACAAACGATTTAGAACAGAGCTTGCTTAAAGAAAGTTAGAA 309
Db 1466 LySGInGInGInGInSeArSpLeuGInGInuArGArGAlaLySGInuLySGIn 1485
QY 310 GAGCAGCAAGGATTTAGAACAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAACA 369
Db 1486 GluGInGInuArGArGAlaLySGInuLySGInGInGInSeArSpLeuGInGIn 1505
QY 370 AGCGATTTAGAACAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAACGATTTA 429
Db 1506 SerArSpLeuGInGInuArGArGAlaLySGInuLySGInGInGInSeArSpLeu 1525
QY 430 GACACAGAGAGACTTGCTTAAAGAAAGTTGCAAGAACAAACGATTTAGAACAGAG 489
Db 1526 GluGInGInuArGArGAlaLySGInuLySGInGInGInSeArSpLeuGInGIn 1545
QY 490 AGACGTGCTTAAAGAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGAGCTGCT 549
Db 1546 ArGArGAlaLySGInuLySGInGInGInSeArSpLeuGInGInuArGArGAla 1565
QY 550 AAAAGAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGAGCT----- 597
Db 1566 LySGInuLySGInGInGInSeArSpLeuGInGInuArGArGAlaLySGInuLy 1585
QY 597 ----- 597
Db 1586 LeuGInGInGInGInSeArSpLeuGInGInuArGArGAlaLySGInuLySGInGIn 1605
QY 598 -----GCTTAAAGAAAGTTGCAAGAGCAGCAAGAA 627
Db 1606 GluGInuArGArSpLeuGInGInuArGArGAlaLySGInuLySGInGInGInuArG 1625
QY 628 GATTTAGAACAAAGAGCTGATACGAAAAAAATTTAGAAAAAGAAAAAGAACATGGA 687
Db 1626 ArSpLeuGInGInuArGArGAlaLySGInuLySGInGInGInuArGArGAlaLySGIn 1645
QY 688 GATATATTAGCAGAGATTTATATGTCGTTTAAAGAAATACAGCTTATAGAACTTCATCA 747
Db 1646 ArSpValLeuAlaGInuArSpLeuGInGInuArGArGAlaLySGInuLySGInGInuAr 1665
QY 748 GAAATATGAGCGATATATATATACACATCAATCTTTTACCTCAGAGAACAGAGGG 807
Db 1666 GluArGInuArGAlaLySGInuLySGInGInGInSeArSpLeuGInGInuArGArGAla 1685
QY 808 AATAGTAGAGATTCAGAGAAATATCTATATAGAAAAAAACAATAGAGAAATCTATTTACA 867
Db 1686 AAnSeArGArSpSeArSeArSeArSeArSeArSeArSeArSeArSeArSeArSeArSeAr 1705
QY 868 ACAATGTTGAAAGAGAGAGATATATACATAAGAAACATCTTGAGAAAAAGAAAGTGT 927
Db 1706 ThrArSpValGInuArGArGAlaLySGInuLySGInGInGInuLySGInuLySGInuLy 1725
QY 928 TCAATATTAACAGAAACAAAGAAAGTAATCTGAGACATCAAAATCATCATTAAGAG 987
Db 1726 SerLeuLySGInGInuArGArGAlaLySGInuLySGInGInGInuLySGInuLySGInuLy 1745
QY 988 ACAATATATTTCTGATGTTATATGATTTTCAATATAGTAAGATGAGATGAATTAAGT 1047
Db 1746 ThrValAAnuLSeArSpValAAnuArSpLeuGInGInSeArSeArSeArSeArSeArSeAr 1765

QY 1048 GCTGAATATGACGATTCATTAATAGATGAGAGAAAGATGATGAGACTTACAGCAATTT 1107
Db 1766 AlaGInuLyArSpArSpSeArSeArSeArSeArSeArSeArSeArSeArSeArSeArSeAr 1785
QY 1108 AAGCTATTTGCAATATGACAAATTTCCAGATGAGAGAAAACTATGAAATTTATTAAGAA 1167
Db 1786 LySGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1805
QY 1168 CTAGAGATTTGATGAGAAAAATGAAAAATTTAGATGATTTAGATGAGAAAGTAAAGAA 1227
Db 1806 LeuGInuArSpLeuLLeGInuLySGInGInuArSpLeuGInuArSpLeuGInuLySGInuLy 1825
QY 1228 TCATCAGAGAAATTAATCTGAGAAAAATTAAGAAAAAGGAAATATGAAAAAACAAG 1287
Db 1826 SerSeArGInuLySGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1845
QY 1288 GATATATATTTTAAACCAATGATTAATAAGTTGATGATGAGCATATTTAAAAAATATTA 1347
Db 1846 ArSpAAnuArSpLeuLySGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 1865
QY 1348 AATGATTAAGCAGCTTAAATAGAAAGAAAGAAATTCATTAATATCATGTTTCATATATTT 1407
Db 1866 AAnuArSpLySGInuArSpLeuGInuLySGInuLySGInuLySGInuLySGInuLySGInuLy 1885
QY 1408 GACGAGACAAATGAAATTTTACAGATCGTGATGAGATTAATCTGAAATATTAATAATAT 1467
Db 1886 ArSpGInuArSpAAnuGInuLLeuGInLLeuAlaLySGInuLySGInuLySGInuLySGInuLy 1905
QY 1468 TTTATGAAACTA 1479
Db 1906 PheMetLySGInuLySGInuLySGInuLySGInuLySGInuLySGInuLySGInuLySGInuLy 1909
RESULT 2
US-09-837-344-31
; Sequence 31, Application US/09837344
; Patent No. US2002041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRULHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:

Db 1 SerapleuglIngluArghalalysegluylseugInglInserapleu 20
Qy 73 GAACAAGATAGACTTGCTAAAGAAAATTACAAAGCCGAAAGCATTTAGAACAAAG 132
Db 21 GluGlnAepArghleuAlalysegluylseugInglInserapleuGlnInglu 40
Qy 133 AGACTTGTAAAGAAAGTTGCAAGAACAAAGCGATCTTGAACAAAGAGAGCTGCT 192
Db 41 ArgaTgAlalysegluylseugInglInserapleuGlnIngluArghAla 60
Qy 193 AAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAAAGAAAG 252
Db 61 LysegluylseugInglInserapleuGlnIngluArghAlalysegluylse 80
Qy 253 TTGCAAGAACAAAGCGATTTAGAACAAAGCTGCTAAAGAAAAGTTACAAAG 312
Db 81 LeugInglInglInserapleuGlnIngluArghAlalysegluylseugInglu 100
Qy 313 CAGCAAGCGATTTGAACAAAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGC 372
Db 101 GlnGlnInserapleuGlnIngluArghAlalysegluylseugInglInglInser 120
Qy 373 GATTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAA 432
Db 121 AepLeuGlnIngluArghAlalysegluylseugInglInglInserapleuGln 140
Qy 433 CAAGAGAGCTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAG 492
Db 141 GlnGlnuArghAlalysegluylseugInglInglInserapleuGlnIngluArgh 160
Qy 493 CGTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAA 552
Db 161 ArgaAlalysegluylseugInglInglInserapleuGlnIngluArghAlalyse 180
Qy 553 GAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAAAGAAAAGTTG 612
Db 181 GluylseugInglInglInserapleuGlnIngluArghAlalysegluylseu 200
Qy 613 CAAGAGCGAAAGAGATTTAGAACAAAGAGCGATTAAGAAAAAATTTAGAAAGA 672
Db 201 GlnGlnInglInglArghleuGlnIngluArghAlalysegluylseuGlnuArgh 220
Qy 673 AAAAGGAGCATGAGATATTTAGAGAGAGATTTATGTCGTTAAATAACACAGCT 732
Db 221 LysegluylseuAlalysegluylseuAlalysegluylseuAlalysegluylseu 240
Qy 733 ATGAGACTTCATCAGAAAATGAACTGGATTTATATATCAATCAATCTTCTTACCT 792
Db 241 lIeGlnleuProserGlnuengluArghAlalysegluylseuProserleuPro 260
Qy 793 CAGGACACAGAGGGAATGTAGAGATTTCCAAAGAAATCTTAAATAAGAAAAACAAAT 852
Db 261 GlnAepAenArghAlalysegluylseuAlalysegluylseuAlalysegluylseu 280
Qy 853 AAGAGATCTATTAACAACAATGTTGAAGAGCGAAGGATATCATTAAGAGACATCTTGA 912
Db 281 ArghuSerleTherThaenValGlnuArghAlalysegluylseuAlalysegluylseu 300
Qy 913 GAAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAAGATTAATCT 960
Db 301 GluylseuAlalysegluylseuAlalysegluylseuAlalysegluylseu 316

RESULT 4
US-09-837-344-20
; Sequence 20, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUBERIN-MARCHAND, Claudine
; DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathie
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-837-344-20
Alignment Scores:
Pred. No.: 4,35e-33 Length: 117
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 24.26% Gaps: 0
DB:
US-09-837-344-41 (1-1482) x US-09-837-344-20 (1-117)
Qy 610 TTGCAAGAGCGACCAAGAGATTTAGAACAAAGAGCGCTGATACGAAAAAATTTAGAA 669
Db 1 LeugInglInglInglArghleuGlnIngluArghAlalysegluylseuGlnu 20
Qy 670 AAAAAAGGAGCATGAGATATTTAGAGAGATTTATATGTCGTTAAATAACCA 729
Db 21 ArghuylseuAlalysegluylseuAlalysegluylseuAlalysegluylseu 40
Qy 730 GCTATAGAACTTCATCAGAAAATGAACTGGATTTATATATCAATCAATCTTCTTTA 789
Db 41 AlalysegluylseuProserGlnuengluArghAlalysegluylseuProserleu 60
Qy 790 CCTCAGACACAGAGGGAATGTAGAGATTTCCAAAGAAATCTTAAATAAGAAAAACA 849
Db 61 ProGlnAepAenArghAlalysegluylseuAlalysegluylseuAlalysegluylseu 80
Qy 850 AATAGAGATCTATTAACAACAATGTTGAAGAGCGAAGGATATCATTAAGAGACATCTT 909
Db 81 AenAlysegluylseuAlalysegluylseuAlalysegluylseuAlalysegluylseu 100
Qy 910 GAAAGAAAAGAGATGTTCAATTAACCAAGAACAAAGAAAGATTAATCT 960

Db 101 GluGluIuLySeLySeApGlySerIleLySeProGluGluIuApLySer 117

RESULT 5

US-09-900-963-20

Sequence 20, Application US/09900963

Publication No. US20030064075A1

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/900,963

FILING DATE: 10-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/098,327

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-045

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 117 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-09-900-963-20

Alignment Scores:

Pred. No.: 4.35e-33 Length: 117

Score: 603.00 Matches: 117

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 24.26% Indels: 0

DB: 3 Gaps: 0

US-09-837-344-41 (1-1482) x US-09-900-963-20 (1-117)

QY 610 TTGCAAGAGCAGCAAGAGATTAGACAAAGAGCTGATACGAAAAAATTAGAA 669

DB 1 LeuGlnGluGlnGlnGlnArgSerLeuGlnGlnArgGlySerThrLeuGlnGln 20

QY 670 AGAAAAAGCAATGATGATATATTAGCAGAGATTATATGTCGTTTGAATAACCA 729

DB 21 ArgLySeLySeGlnHISGlyASeRleLeuAlaGluASeRleuLyGlyArgLeuGlnIlePro 40

QY 730 GCTATGAACTTCCTCATGCAAAAAATGAACTGATATTATATCAATCATCTTCTTTA 789

Db 41 AlaIleGluLeuProSerGluAenGluAArgGlyTyrTyrIleProHISGlnSerSerLeu 60

QY 790 CCTCAGACCAACAGAGCAATAGTAGATTCACAGAAATATCTTATAGAAAAACA 849

DB 61 ProGlnASePasnArgLyASeNSeRArgASeRLeLyGluIleSerIleIleGluLyThr 80

QY 850 AATAGAGATCTATTCAACCAATGTTGAAGACGAGGATATACATAAAGCAATCTT 909

DB 81 AsnArgGluSerIleHISThrAsnValGluGlyArgArgAspIleHISLySGlyHISLeu 100

QY 910 GAAGAAAAGAAAGATGTTCAATTAACCAAGCAAAAGAAAGATTAATCT 960

DB 101 GluGluIuLySeLySeApGlySerIleLySeProGluGluIuApLySer 117

RESULT 6

US-09-837-344-19

Sequence 19, Application US/09837344

Patent No. US20020041882A1

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/837,344

FILING DATE: 19-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/462,625

FILING DATE: <Unknown>

APPLICATION NUMBER: FR 91 01286

FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-09-837-344-19

Alignment Scores:

Pred. No.: 9.54e-30 Length: 107

Score: 554.00 Matches: 107

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 22.28% Indels: 0

DB: 3 Gaps: 0

```

US-09-837-344-41 (1-1482) X US-09-837-344-19 (1-107)
QY 640 AGAAGAGCTGATACGAAAAAAATTTAGAAAAGAAAAAGAACACTGGAGATATTATTACGA 639
Db 1 ArglysAlaAspTrpTyrLysAlaSerLeuGlnArgLysLysGluHisGlyAspIleLeuAla 20
QY 700 GAGGATTTATATGGCGCTTTAGAAAATACACGCTTTAGAACTTCATCAGAAAATGAACGT 759
Db 21 GlnAspLeuTyrGlyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArg 40
QY 760 GGATATTATATACACATCAATCTTCTTTACCTCAGACAAACAGAGGAAATAGTAGAGT 819
Db 41 GlyTyrTyrIleProHisLeuGlnSerLeuProGlnAspAsnArgGlyAsnSerArgAsp 60
QY 820 TCCAAAGAAATATCTTATATAGAAAAAACAATAGAGATCTATTACAACAAACTTGAA 879
Db 61 SerTysGluIleSerIleIleGluLysThrAsnArgLysSerIleThrAsnValGlu 80
QY 880 GGAGCAAGCGATATATCATPAAGAGACATCTTGACAGAAAAAGAAAGATGGTTCATPAACCA 939
Db 81 GlyArgGAspIleHisLysGlyHisIleuGluGluLysLysAspGlySerIleLysPro 100
QY 940 GAACAAAAAGAGATAAATCT 960
Db 101 GluGlnLysGluAspLysSer 107

RESULT 7
US-09-900-963-19
; Sequence 19, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRULIHE, Pierree
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Unl-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13864

```

```

; PUBLICACION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-900-963-19

Alignment Scores:
Pred. No.: 9.54e-30 Length: 107
Score: 554.00 Matches: 107
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.28% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-41 (1-1482) x US-09-900-963-19 (1-107)

OY 640 AGAAGAGCTGATACGAAAAAATTTAGAAAGAAAAAGAAACATGAGATATATTAGCA 639
Db 1 ArgysaiaaaprtlnlyelaysanleuiliuArglySLySGlInHIGlYApriIleuila 20
OY 700 GAGGATTTATATNGCGCTTTGAATATACAGCTATAGAACTTCACATCAGAAATAGAACGT 759
Db 21 GlusapleuryrGYLxArgLeuglInleProalalIeGluIenProserGIuansgluArg 40
OY 760 GGATATTTATATACACATCAATCTCTTTACTCTCAGACAACAGAGGAGATAGAGAT 819
Db 41 GLyrrYrrYrIleProHieGlnserSerleuProGlnApsrArlxGlyAmsrArGAp 60
OY 820 TCACAGGAATATCTATATATGAAAAAACAATAGAAATCTATTATCAACAAATGTTGAA 879
Db 61 SerlysgIuIleSerIleIeGluIuYthrAanArGJuseRilethrAanValGIu 80
OY 880 GGACGAGCGATATACATTAAGAGACATCTTGAAAGAAAGATGCGTTCATTAACCA 939
Db 81 GLyArGArGArpIleHieSLySGlYHIsleuglInGluLySLyApselYserIleSLyPro 100
OY 940 GAACAAAAGAAGATAATCT 960
Db 101 GluGlnlySGluApslySer 107

RESULT 8
US-10-732-923-3319
; Sequence 3319, Application US//10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3319
; LENGTH: 1003
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-10-732-923-3319

Alignment Scores:
Pred. No.: 5.35e-18 Length: 1003
Score: 381.50 Matches: 101
Percent Similarity: 53.41% Conservative: 126
Best Local Similarity: 23.76% Mismatches: 105
Query Match: 15.35% Indels: 93
DB: 5 Gaps: 15

US-09-837-344-41 (1-1482) x US-10-732-923-3319 (1-1003)

OY 4 GAACACAAAGCGACTCTAGAACACAGAGACGCTGTCAAGAAAAGTTGCAAGAACAA 63
Db 440 GInGInGInGInuProGInGInGInGInuProGInGInGInuProleuGInGInuProGln 459
OY 64 AGCGATTTAGAACAGATGACTTGCTATAAGAAAAGTTACAAAGACGACGAAAGCGATT 123
Db 123 AGCGATTTAGAACAGATGACTTGCTATAAGAAAAGTTACAAAGACGACGAAAGCGATT 123

```

[illegible]

QY	1195	AATTAGATGATTTAAGATGAA-----CGAATTAGAAAAATCATCAGAAGATTA	1242
D8	754	GluLeuGIngluValGluglIngluIngluInclValGluglIngluIngluIngluthr	773
QY	1243	TCTGAAGAAAAATA 1257	
D8	774	ValGlugluProlle 778	
<hr/>			
RESULT 9			
US-10-732-923-3359			
; Sequence 3359, Application US/10732923			
; Publication No. US20050108791A1			
; GENERAL INFORMATION:			
; APPLICANT: Edgeton, Michael D			
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES			
; FILE REFERENCE: 38-15 (52796) C			
; CURRENT APPLICATION NUMBER: US/10/732,923			
; PRIORITY FILING DATE: 2003-12-10			
; PRIOR APPLICATION NUMBER: 10/310,154			
; PRIOR FILING DATE: 2002-12-04			
; NUMBER OF SEQ ID NOS: 24149			
; SEQ ID NO 3359			
; LENGTH: 1407			
; TYPE: PRF			
; ORGANISM: Oryctolagus cuniculus			
US-10-732-923-3359			
<hr/>			
Alignment Scores:			
Pred. No.:	6,23e-18	Length:	1407
Score:	380.50	Matches:	127
Percent Similarity:	45.84%	Conservative:	132
Best Local Similarity:	22.48%	Mismatches:	167
Query Match:	15.31%	Indels:	139
DB:	5	Gaps:	20
<hr/>			
US-09-837-344-41 (1-1482) x US-10-732-923-3359 (1-1407)			
QY	1	CAAGAACACAAAGCGATCTA-----GAACAAGAGACGTGCTAAAGAAAGTGG	51
D8	140	GIngluIngluArGLuLeuAlaGlugluIngluIngluIngluIngluValvArGLuArGphe	159
QY	52	CAAGAA-----CAACAAGCGATTTAAGAACAGATGACTTGCTAAAGAAAG	99
D8	160	GluginhIstYrSerArGlnTrYAraAspLySgluIn---ArgLeuGInArGInglu	178
QY	100	TTACAAGAGCAGCAAAGCGATTTTAGAACAMAGAGACTTGCTAAAGAAAGTGCAGAA	159
D8	179	LeuGIngluArGLaAla-----GlugluIngluInleuArGLyGlyArGLyArGLap	196
QY	160	CAACAAGCGATCTAAGAACAGAG-----AGCGTGTAAAGAAAAGTTCCAGAA	210
D8	197	AlaGlugluPheIIeGlugluIngluInleuArGLyArGLyGIngluInleuLysArg	216
QY	211	CAACAAGCGATTTAGAACAGAGAGACGTGCTAAAGAAAG-----TTG	255
D8	217	GIUeuArGLuGluIngluIngluInArGLyArGLuArgGLuInhIsGLuArGLAleu	236
QY	256	CAAGAACAACAAGCGATTTAGAACAMAGATGACTTGCTAAAGAAAGTTACAAAGACG	315
D8	237	GIngluIngluIngluInleuLeuArGLInArGLyArGLTPArGLuInProArGLuIn	256
QY	316	CAAGCGATTTAGAACAGAGAGACGTGCTAAAGAAAGTTGCCAAGAACAAAGCGAT	375
D8	257	GIn-----GInleuArGLyArgGLuLeuGluInGluIleArGLuArGLuInArg	272
QY	376	TTAGAACAGAGAGACGTGCTAAAGAAAGTTGCCAAGAACAAAGCGATTTAGAACAA	435
D8	273	LeuGluIngluIngluArGLyArgGLuIngluInleuArGLyArgGLuInArg--LeuGluIn	291
QY	436	GAG-----AGACTTCTAAAGAAAGTTGCCAAGAACAAAGCGAT	474
D8	292	GIUGluArGLyArgGLuInleuArGLyArgGLuLeuGluInGluIleArGLuArGLuIn	311

Oy	475	GATTATGAGCAACAAGAGACGCGCTAAACAAAGTTTGCAGAAACAACAAGCGATTTGAA	534
Oy	312	ArgLeuGluGlnGluArgGluArgGluGlnArgLeuGlnGlnGluGluArgGluGln	331
Oy	535	CAAGAGACGCTGCTAAAGAAAGTTGGACAAACAACAAGCGATTTGAA	594
Oy	332	GlnLeuIysArgGluLeuGlnGluIleArgGlnGluArgGlnArgGlnArgGlnGlu	351
Oy	595	CGTGTAAAGAAAGTTCTCCAAAGACGACCAAAAGATTTGAAACAAGAAGCGTGAACG	654
Oy	352	ArgArgGlnGlnLeuLeuAlaGlnIleValArg	362
Oy	655	AAAAAAATTTAGAAAGAAAGAAACATGACGATATTATAGCAGAGATTTATGCT	714
Oy	363	-----GluGlnAlaArgGluArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	378
Oy	715	CGTTTAGAATAACGACTATAGAACTTCATCAGAAAATGAACCTGCATTTATTATCA	774
Oy	379	GlnLeuGlu-----SerGlnAlaGlnIleAlaArgGlnSerIleValTyrSerArgPro	395
Oy	775	-----CATCATCTTCTTTTACCTGAGGACAACAAGGGAATATGATAGATTTCCAG	825
Oy	396	ArgArgGlnGlnGluGlnSerLeuArgGlnAspGlnIleArgArgGlnArgGlnGlnArg	415
Oy	826	GAAATATCTATATATGAAAAAAACAATGA	855
Oy	416	GluArgGlnLeuGlnGlnGlnAlaArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	435
Oy	856	GAATCTATTACAAACAATCTTCAGAGACGAAAGGATATACATTAAGACATCTT	909
Oy	436	GluArgArgArgGlnArgLeuSerAlaArgProSerLeuArgGlnGlnGlnGlnGlnGlnGln	455
Oy	910	GAAGAAAGAAAGATGCTTCAATAAA-----CCAGAACAAAAAGAGATAA	957
Oy	456	GlnGluIleArgGlnGlnGlnGlnArgPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	475
Oy	958	TCTGCTGACATACAA-----	972
Oy	476	ArgGlnGlnLeuGlnPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	495
Oy	972	-----	972
Oy	496	GlnLeuGlnGlnGlnAspSerPheGlnIleAspArgGluArgArgArgGlnGlnGlnGln	515
Oy	973	-----AATCAT	978
Oy	516	GlnArgProGlnGlnThrTyrArgTyrPrgGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	535
Oy	979	ACATT-----GAGCAGCTAAATATTCTGATGTTAAAGATTTCGAATA	1023
Oy	536	ThrLeuTyrAlaLysPProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	555
Oy	1024	AGTAAGTTAGAGAGAAATAAGTCTGCAAT--GACGATTCATTATAGATGACAGAA	1080
Oy	556	GluIysArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	575
Oy	1081	GAAGATGATGAAGACTTAGACGAATTTAAAGCTATTGTGCATATGACAAATTTCCAGAT	1140
Oy	576	GluAspGluLysArg-----ArgArg	582
Oy	1141	GAAGAAACATGAGAAATTATTAAGAACTGAGAAAGTTTGATGAGAAATAAGAAATTTA	1200
Oy	583	GlnGlnIleArgGlnArgGlnTyrArgGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	602
Oy	603	AspArgGlyLeuArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	622
Oy	1261	AAAGGAAAGAAATATGAAAAACAAGATAAT-----TTTAAACCAATGATATA	1314
Oy	623	ArgGlnGlnGlnGlnArgGlnArgLeuArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	642

Qy	1315	AGTTTGATGATGAGCATATTAATAAAATATTAATAATGATTAAGCAAGTTAATAAGAAAG	1374
Db	643	Glueuargingluarvgluarvlyleuargluugluuglnleuuarargvlu	662
Qy	1375	GAATAATTCAATMAA	1389
Db	663	Gluingluuearvg	667
RESULT 10			
US-10-732-923-3358			
/ Sequence 3358, Application US/10732923			
/ Publication No. US200501087911			
/ GENERAL INFORMATION:			
/ APPLICANT: Edgerton, Michael D			
/ TITLE OF INVENTION: "TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES			
/ FILE REFERENCE: 38-15(52796)C			
/ CURRENT APPLICATION NUMBER: US/10/732,923			
/ CURRENT FILING DATE: 2003-12-10			
/ PRIOR APPLICATION NUMBER: 10/310,154			
/ PRIOR FILING DATE: 2002-12-04			
/ NUMBER OF SEQ ID NOS: 24149			
/ SEQ ID NO 3358			
/ LENGTH: 1549			
/ TYPE: PRT			
/ ORGANISM: Ovis aries			
US-10-732-923-3358			
Alignment Scores:			
Pred. No.: 6,23e-18 Length: 1549			
Score: 380.50 Matches: 123			
Percent Similarity: 43.59% Conservative: 139			
Best Local Similarity: 22.13% Mismatches: 194			
Query Match: 15.31% Indels: 145			
DB: 5 Gaps: 18			
US-09-837-344-41 (1-1482) x US-10-732-923-3358 (1-1549)			
Qy	1	CAAGAACAAACAAAGCGATCTAGAACAAAGAGACGCTGCTAAAGAAAGTTGCAGAACAA	60
Db	197	Gluarvglnarvleuglnvalgluglnlgnlnarvglnarvglyleuargvluarvg	216
Qy	61	CAAAAGCATTTA-----GAACAGTTAGCTTCTAAAGAAAGTTACAA	105
Db	217	GlnglgluvalgluInleuglnlyearvglnurthrglngluInleuglnarvglnurgleuglu	236
Qy	106	GAGCAGCAAAAGCGATTAGAACAAAGAGACGCTGCTAAAGAAAGTTGCAGAACAA	165
Db	237	Gluglglnglnleuglnlyearvglnlyearvglnleuglnleuglnleuglnleugln	256
Qy	166	AGCGATCTAGAACAAAGAGACGCTGCTAAAGAAAGTTGCAGAACAAACAAAGCGATTTA	225
Db	257	Argarvglnuglnleuuarargvlyearvglnleuglnuarvgargvlglnleuuarvgln	276
Qy	226	GAACAAAGAGACGCTGCTAAAGAAAGTTGCAGAACAA-----	264
Db	277	GluglglnglnualarhrcnglnlgnlulleservglnuarvglyluserargThrserarg	296
Qy	265	-----CAACGCGATTAGCAAGATAGCTGCTGCTAAAGAAAG-----	303
Db	297	CyvgIntTpglnleuglnusercglualaaspalaarvglnarvlyearvalTysersargPro	316
Qy	304	---TTACAAAGACCAAAAGC-----GATTAGAACAAAG	336
Db	317	Hlsarvglnuglnlgnlserargvrgnglnuglnleuueuglnarvglnglnlgn	336
Qy	337	AGACGTCTAAAGAAAGTTGCAGAACAAACAAAGCGAT-----TTAGAA	381
Db	337	Glnglnlleservglnvalglnserleuglngluarvglnglyarvglnarvglyuys	356
Qy	382	CAAGAGAACGCT-----	393
Db	357	GlnglgluInarvTyarspcInlansrTpargrTpglnleuglnleuglnusercglnarvg	376

QY 394 -----GCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAG 438
Db 377 ArgTyrThrLeuTyrAlaIleProAlaGlnArgGlnGlnValArgGlnGlnGlnLeu 396
QY 439 AGACTTCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGA----- 492
Db 397 ArgLeuIleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 416
QY 493 -----CGTCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAA 534
Db 417 ArgGlnValGlnLeuGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 436
QY 535 CAAGAGAGAGCTGCTTAAGAAAGTTGCAAGAA----- 567
Db 437 ArgGlnGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 456
QY 568 -----CAACAAGCGATTAGAACAAAGAGAGCGTCTAAAGAAAGTTGCAAG 618
Db 457 GlnLeuTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 476
QY 619 CAGCAAGAGATTAGAA-----CAAGAAGGCT 648
Db 477 GlnIleGlnTyrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 496
QY 649 GATCGCAAAAAAATTAGAAAGAAAAAG-----GAACATGAGATATATTAGCA 699
Db 497 ArgGlnIleuArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 516
QY 700 GAGATTATATGCGCTTAAAGAAATACAGCTATAGAACTTCAATCAGAAAGAAAGCT 759
Db 517 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 534
QY 760 GGATATATAT-----ATACCATCATCTTCTTAACTCAGAGCAAGAGG 807
Db 535 GlnTyrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 554
QY 808 AATAGTAGAGATTCAGAGAAATCTATATAGAAAAAACAAATAGAGAA----- 858
Db 555 LysArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 574
QY 859 -----TCTATTACAACAAAATGTTGCAAGAGAGGATATACATTAAGAAAGCTTGA 912
Db 575 GlnLeuGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 594
QY 913 GAA-----AGAAAGATGCTTCAATAAACCAAGAAAGAAAGAAAGATTCGCT 963
Db 595 LysValGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 614
QY 964 GACATACAAATCATCATTTAGAGACAGTAAATTTCTGATGTTAATGATTTCA--- 1020
Db 615 GlnArgGlnArgGlnTyrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 634
QY 1021 -----ATTAGT 1026
Db 635 GlnGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 654
QY 1027 AAGTATGAGATGAATAAGTCTGCAATATGACATTCATTAATGATGAAGAAAGAT 1086
Db 655 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 672
QY 1087 GATGAAGCTTAGAGCAATTTAAGCTTGTGCAATATGACAAATTTCCAAAGTAGAGAA 1146
Db 673 LysArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 692
QY 1147 AACATTA-----GGAATTTAAAGAA 1167
Db 693 GlnLeuGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 712
QY 1168 CTAGAAGATTGATAGAGAAAGAAATTTAGATGATTTAGATGAAGAAATAGAGAA 1227
Db 713 LysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 732
QY 1228 TCATCAGAGAA-----TTATCTGAAGAAAAATA 1257

Db 733 GlnArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 752
QY 1258 AAAAAAGAGAAATATGCAAGAAAGATTAATTTAAACCAATGATTAAGT 1317
Db 753 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 772
QY 1318 TTGTATGATGAGCATATTTAAATAATTAAGATAGAGGTTAATAGAGAAAGAA 1377
Db 773 GlnTyrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 792
QY 1378 AAA 1380
Db 793 Arg 793
RESULT 11
US-10-450-763-47660
Sequence 47660, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OR INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIORITY APPLICATION NUMBER: PCT/US01/08631
PRIORITY FILING DATE: 2001-03-30
PRIORITY APPLICATION NUMBER: 09/540,217
PRIORITY FILING DATE: 2000-03-31
PRIORITY APPLICATION NUMBER: 09/649,167
PRIORITY FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 47660
LENGTH: 419
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (362)..(408)
OTHER INFORMATION: Vertebrate metallochioneins proteins domain identified by
OTHER INFORMATION: EMATRIX, accession number BI00203, p-value=2.484e-11, raw score
OTHER INFORMATION: 13.94
NAME/KEY: misc_feature
LOCATION: (1)..(419)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-47660
Alignment Scores:
Pred. No.: 1,19e-17 Length: 419
Score: 376.50 Matches: 149
Percent Similarity: 44.34% Conservative: 35
Best Local Similarity: 35.90% Mismatches: 76
Query Match: 14.26% Indels: 155
Gaps: 25
US-09-837-344-41 (1-1482) x US-10-450-763-47660 (1-419)
QY 1217 CCTTCATTAATCATCTAAATTTTCAATTTTCTATCAAACTTCTAGTCTTTATTA 1158
Db 136 ProAlaProAlaSerSerProCySer---ArgSerLeuSerSerSerArgSerLeu--- 153
QY 1157 ATTCTATGTTTCTTCATCTTGAGAA-----TTGCATATTGCACA 1116
Db 154 -----SerSerSerCyArgValSerCySerHisSerLeuSerSerCySer 169
QY 1115 ATAGCTTAATTTGCTTAAGTCTTCATCATCTTCTTCATCATATTAATGAATGCTCA 1056
Db 170 ArgSerLeuSerCySerArgSerSerSerThSerArgSerSerThSerSer 189
QY 1055 TATTCAGACTTATTTCTCATCTTACTTACTTATTTGAAGAAATCAATTAACATAGAGAAATA 996


```
QY 391 -----CGTGTAAAGAAAAGTTGCAAGAACAAAGCCGATTGAGAACAGAGAGA--- 441
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 756 GlnGluArgGluGluGluGlnLeuArgAspArgProSerArgArgGluGlnGluArgHis 775
QY 442 -----CTTGCTAAAGAAAAGTTGCAAGAACAAAGCCGATTGAGAACAGAGAGA--- 492
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 776 GlnGluArgGluGluGlnLeuArgAspArgProSerArgArgGluGlnGluArgHis 795
QY 493 -----CGTGTAAAGAAAAGTTGCAAGAACAAAGCCGATTGAGAACAGAGAGA--- 543
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 796 GlnGluArgGluGluGlnLeuArgAspArgProSerArgArgGluGlnGluArgHis 815
QY 544 -----CGTGTAAAGAAAAGTTGCAAGAACAAAGCCGATTGAGAACAGAGAGCGT 597
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 816 LeuGluArgGluGluGlnLeuArgAspArgProSerArgArgGluGlnGluArgHis 835
QY 598 GCTAAAGAAAAGTTGCAAGAGCAGCAAGATTTAGAACAAAGGCGTGATACGAAA 657
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 836 ---GlnGluArgGluGluGlnLeuArgAspArgProSerArgArg---GluGlnGlu 853
QY 658 AAAAATTTAGAAAAGAAAAGAAACATGAGATATATTAGCAGAGATTTATATGCTCGT 717
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 854 ArgArgLeuGluArgGluGluGlnGln-----LeuArgAspArgSerPheArgArg 870
QY 718 TTAGAAATACCAGCTATAGAACTTCATCAGAAAATGACGCGATTTATATCCACAT 777
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 871 GluGln-----GluLeuArgArgAspArgPheHisGluGluGluArg 886
QY 778 CAATCTTCTTACTCAGGACAAACAGAGGGAATGTAAGAGAT----- 819
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 887 ArgGluGluLeuGluGluGlnGlnArgGlyGlnGluArgAspArgGluGlnGlu 906
QY 819 ----- 819
Db 907 GlnLeuArgGluGlnArgGluGluGlnArgArgArgGlnGlnGlyAspArgGlyLeu 926
QY 820 TCCAGAGAAATATCTATATATGAAAAAACAATAGAGAA-----TCTATT 864
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 927 HisArgGluLeuGlnGluValArgGlnGluGlnGluArgLeuArgAspArgGlyLeu 946
QY 865 ACAACAATGCTTGAAGACGAGAGATTTACATAAGACACTCTTGAAGAAAAGAAAGAT 924
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 947 ArgArgGluGlnGlnLeuArgArgAsp---ArgGlySerPheHisGluGluGluArg-- 964
QY 925 GGTCAATTAATAACAGAAACAAGAAAGATTAATCTGCTGACATACAAAATCAT----- 978
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 965 -----ArgHisGluGluPheGluGlnGlySerGlnLeuArgLeuGlnGluProAspArg 981
QY 979 -----ACATTAGAGACAGTAATATTTCTGATGTTAATGAT 1014
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 982 ArgPheArgArgGluGlnGlnLeuArgGlnGlnGlyValGluGlnGluArgLeuArgAsp 1001
QY 1015 TTTCAATTAAGTAAGTATGAGATGAAATAAGTCTGCTGAATATGACGAT----- 1062
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1002 SerGlyHisLeuArg---GluGlnGlnLeuArgArgGluArgGluGlnGluArg 1020
QY 1063 -----TCATTATAGATGAGAGAAAGAT 1086
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1021 AspArgGlyHisLeuArgArgAspGlnGlnLeuArgGlnGlnGlyLeuGlnGlnLeuArg 1040
QY 1087 GATGAGAGCTTAGAC---GAATTTAAGCCTATTGCAATATGCAATTTTCCAGATGAA 1143
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1041 ArgGlnGlnLeuArgAspArgGlyPheArgGluGlnGlnGlnLeuArgGlnGlnGlu 1060
QY 1144 GAAAAACATA-----GGAATTTATTAAGAACTAGAAAGTTGATAGAG----- 1185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1061 GlnArgLeuArgAspArgGlyHisLeuArgArgGluGlnGlnLeuArgArgGluGlnGlu 1080
QY 1186 AAAAATTAATTTAGATGATTTAGAGAAATGAGAAATATCATCAGCAAGAAATTAATCT 1245
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1081 ArgArgGluGlnGlnPheArgArgGluGlnGlnLeuArgArgGluGlnGlnPheArgArg 1100
```

```
QY 1246 GAGAAAAAATAAAGAAAGAAATATGAAAAAACAAGGATATATTAAACCA 1305
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1101 GlnGlnGlnLeuArgGlnGlnGluGlnGluArgLeuArgAspArgGlyHisLeuArgArg 1120
QY 1306 AATGAT-----AAAAGTTGTATGATGACATATTAATAAATATTAATTAAG 1356
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1121 AspGlnGlnLeuArgGlnGlnGlyLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1140
QY 1357 CAGCTTAATTAAGAAAAAGAA 1377
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1141 LysPheArgGluGlnGlnGln 1147

RESULT 13
US-101-487-58
? Sequence 58, Application US/10101487
? Publication No. US20020169125A1
? GENERAL INFORMATION:
? APPLICANT: LEUNG, DAVID W.
? APPLICANT: BERGMAN, PHILIP A.
? APPLICANT: LOFOUST, ALAN A.
? APPLICANT: PIETZ, GREGORY E.
? APPLICANT: TOMPKINS, CHRISTOPHER K.
? APPLICANT: WAGGONER JR., DAVID W.
? TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES
? TITLE OF INVENTION: THERMOF
? FILE REFERENCE: 077319/0329
? CURRENT APPLICATION NUMBER: US/10/101,487
? PRIOR FILING DATE: 2002-03-20
? PRIORITY APPLICATION NUMBER: 60/277,705
? PRIOR FILING DATE: 2001-03-21
? NUMBER OF SEQ ID NOS: 116
? SOFTWARE: Patentn Ver. 2.1
? SEQ ID NO 58
? LENGTH: 350
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
US-101-487-58

Alignment Scores:
Pred. No.: 1,59e-16 Length: 350
Score: 360.00 Matches: 78
Percent Similarity: 50.24% Conservative: 132
Best Local Similarity: 18.66% Mismatches: 140
Query Match: 14.48% Indels: 68
DB: 4 Gaps: 2

US-09-837-344-41 (1-1482) x US-10-101-487-58 (1-350)
QY 1 CAAAGAACAAAGCGATCTAGAACAGAGAGCGCTAAAGAAAAGTTGCAAGAACAA 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 20
QY 61 CAAAGCGATTTAGAACAAAGTATAGACTTCTAAAGAAAAGTTCAAGAGCGCAAGCGAT 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 21 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
QY 121 TTAGAACAGAGAGACTTCTTAAGAAAAGTTGCAAGAACAAAGCGATCTTAGAACAA 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 41 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
QY 181 GAGAGAGCGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGCGT 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
QY 241 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGTATAGACTTCTTAAGAA 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
QY 301 AAGTTACAGAGCGCAAGAGCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCA 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 101 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
```



```
QY 1138 GATGAGAAACATAGAAATTATTAAGAACTAGAGATTGTAGAGAAAAATGAA--- 1194
Db 877 GIUGIUGIn-----GIUleUGIUGIuValGIUGIUGInGIUGIn 890
QY 1195 AATTAGATGATTAGATGAGAAATAGAAAAATCATCAGAAATTAATCTGAGAAAAA 1254
Db 891 GIUleUGIUGIuValGIUGIUGInGIUGInGIUGInGIUGInGIUGInGIUGIn 910
QY 1255 ATAAAAAAGAAAGAAATATGAAAAACAAAGATATTAATTTAAACCAATGATMAA 1314
Db 911 GIUGInGIUleUGIUGIuValGIUGIUGInGIUGInGIUGInGIUGInGIUGIn 930
QY 1315 AGTTGTATGATGAGCATATT 1335
Db 931 GIUthrValGIUGInProIle 937
```

Search completed: December 7, 2005, 07:15:55
Job time : 154 secs

QY	61	CAAGCGCATTTA-----GAAACAAGATAGACTTGCTCTAAAGAAAAGTTGACAAAGACAGCAAGC	117
Db	760	AlatrrrgltuleuSerLysrGlnleuGluMetGluLysGluLevalSerSerThrArgLeu	779
QY	118	GATTTAGAACAAGAGAGACTTGCTTAAAGAAAAGTTGCACAGAACAAACGACATCTAGAA	177
Db	780	AspleuGlnLysrLysSerGluAlaLeuGluSerIleLysrGlnLysrLeuThrLysrGlnLys	799
QY	178	CAAGGAGACGT-----GCTAAAGAAAAGTTGCCA	207
Db	800	GluGluLysrLysrIleLeuLysrGlnAspPheGluThrLeuSerGlnGluThrLysrIleGln	819
QY	208	GAACACAA-----AGCGATTTAGAACAAAGACGCT	240
Db	820	HtseGluGluLeuAsnAsnAsnArgIleGlnThrThrValThrGluLeuGlnLysrValLysMet	839
QY	241	GCTTAAGAAAAGTTGCCAAGAACAAACGACATTTAGAACAGATPAGACTTGCTTAA---	297
Db	840	GluLysrGluAlaLeuMetThrGluLeuSerThrVal---LysAspLysrLysSerLysVal	858
QY	298	---GAAAAGTTACAAAGACGCAAAAGCATTTAGACAAAG---	336
Db	859	SerAspSerLeuLysrAsnSerLysrSerIlePheGluLysrGluAsnGlnLysrGlyValAla	878
QY	337	-----AGACGTGCTTAAAGAAAAGTTGCCAAGAACAA---	366
Db	879	AlaIleLeuAspLeuGlnLysrThrCysrLysrGluLeuLysrHtseGlnLeuGlnValGlnMet	898
QY	367	CAAGCGCATTTAGAACAAAGACAGCGTCTTAAAGAAAAGTTGCCAAGAACAAACGACAT	426
Db	899	GluAsnThrLeuLysrGluGlnLysrGluLeuLysrLysrSerLeuGluLysrGluLysrGluAla	918
QY	427	TTAGAACAAAGAGAGACTTGCTTAAAGAAAAGTTGCCAAGAACAA-----CAAGC	474
Db	919	SerHtseGlnLeuLysrGluLeuAsnSerMetGlnGlnGlnLeuIleGlnAlaGlnAsn	938
QY	475	GATTTAGAACAAAGACAGCGTCTTAAAGAAAAGTTGCCAAGAACAAACGACATTTAGAA	534
Db	939	ThrLeuLysrGlnAsnGlnLysrGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnLysr	958
QY	535	CAAGGAGACGCTTAAAGAAAAG---TTGCACAAACAAACAGCGATTTAGAA-----	585
Db	959	GlnSerSerGluGlnLysrLysrGlnIleGluAlaLeuGlnGlnGlnLysrIleAla	978
QY	586	-----CAAGAGAGACGCTTAAAGAAAAGTTGCCAAGACAGCAAGAGATTTAGAACAA	639
Db	979	ValleuGlnLysrThrGluLeuGlnLysrLysrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	998
QY	640	AGGAAGCGTGATCCGAAAAA-----AATTAGAAAGAAAAAG	678
Db	999	GluLeuAlaAlaGlnLysrGlnLysrIleSerValleuGlnAsnAsnAsnArgLysrLysrGln	1018
QY	679	GAACATGAGATATATATTTAGACAGAGATTTATATGCTCGTTTAGAAATCCACGATATAGAA	738
Db	1019	GluThrThrPheLysrGlnLeuGlnSerAspPheThrGlyArg-----GluSerGln	1034
QY	739	CTTCCATCAGAAAATGAACGTCGATATATATATATACCAATCAATCTTCTTACCTCAGAC	798
Db	1035	LeuLeuAlaThrArgGlnAspLeuLysrSerValGluGluLysrLeuSerLeuAlaGlnGln	1054
QY	799	AACAGAGGAATAGTATAGATTTCCAAAGAAATATCTATATATAGAAAACAAATATAGAA	858
Db	1055	AspLeuIleSerAsnAsnArgAsn-----GlnIleGlyAsnGlnAsnLysrLeuIleGlnLys	1072
QY	859	TCTATTACCAAAATGTTCAAGACGAAGCAAGATATATCATATTAAGACATCTTGAA---GAA	915
Db	1073	LeuLysrThrAlaLysrAlaThrLeuGlnGlnAspSerAlaLysrLysrGluGlnGlnGlnGln	1092
QY	916	AAGAAAGATGCTTCAATAAACCAGAACAAAAAAGAAAT-----AAATCGTCGACATA	969
Db	1093	GluArgCysrLysrLysrLeuGlnAspIleGlnLysrGluLysrSerLeuLysrGluLysrGlnLysr	1112

OY		970 CAAATCTACATTAGGAGCAGTAATAATTATTTCGTGGTGAATTGATTTCCAAATAGAATAG	1029
Dd		1113 ValANGluLySSerLYLeuAlGLnULeLUguLuLleLyS-----CyArg	1128
OY		1030 TATGAGCATTAATAAATAGTGCTGAATATGCATTCAATTATAGATGAAGAAAGATGAT	1089
Dd		1129 GlnGUlySGlUILLethrLYs-----LeuANGluLUleuLySer	1142
OY		1090 GAAGACTTAGACGAATTTAAGCSCTATTGTGCCAATATGCACAATTTCCAAAGATGMAAGAAA	1149
Dd		1143 HisLYLeuGuLSerILeLYeGLUlIethr-----AsnLeuLYsaSPALalYsgLn	1155
OY		1150 ATAGGAATTTAATAAACCTAGAA-----GATTTC	1179
Dd		1160 LeuEUllEGInGUlySaRgaNGInGlInIlEeuLYsaERGInValLYSGlUGlUGlUGlU	1179
OY		1180 ATAGAGAAAAATGAAAAATTTAGATGATTTAGATGAAGGAATAGAAAAATCATCGAAGAA	1239
Dd		1180 GlUGInGUlySaRgaNGInGlInIlEeuLYsaERGInValLYSGlUGlUGlUGlUGlU	1199
OY		1240 TTATTGGAAGAAAAATTAATAAAAAAAAAAGAAATATGAA-----AAAACAAAGAT	1299
Dd		1200 LeuLYLYSGlUPheILEgLUlsGUlAlalySeuHisSERglULelYsgLUlySGlU	1215
OY		1291 AATAATTTAAACCAATGATGATA-----AGTTTGAT	1322
Dd		1220 ValGIyMeLyuLYenHIGlUGlUsnGUlAlalySeuthrmecGLnlEthrLAleu	1235
OY		1324 GATGAGCATATTAATAAATAATTAATAAATGATAGCACGTTAATAAGAAAAAGAAAAATTC	138
Dd		1240 ANGluANleuGLUTThrVallYeLGluLPrgInsERserGINarGaRValsERglU	1253
OY		1384 ATAATAATCATGTTTCATATATTGACGAGAGACAATGAAATTTTACAGATGCTGATGAG	1443
Dd		1260 LeuGUlyvSGIntThrAPraPLeuArGYlgIUlleaLYvalLeuGUlAlathrVaIgLn	1275
OY		1444 TTAATCTGAAGAT	1455
Dd		1280 AsnANGlnADPr	1283
		RESULT 2	
	/ Sequence 2, Application US/11060914		
	/ Publication No. US20050261184A1		
	GENERAL INFORMATION:		
	APPLICANT: Kaufman, Paul L.		
	APPLICANT: Geisger, Benjamin		
	APPLICANT: Berishadsky, Alexander		
	APPLICANT: Borras, Teresa		
	TITLE OF INVENTION: Method for Treating Gliaucoma		
	FILE REFERENCE: 960296.00128		
	CURRENT APPLICATION NUMBER: US/11/060,914		
	CURRENT FILING DATE: 2005-02-18		
	PRIOR APPLICATION NUMBER: US 60/545,722		
	PRIOR FILING DATE: 2004-02-18		
	PRIOR APPLICATION NUMBER: US 60/545,723		
	PRIOR FILING DATE: 2004-02-18		
	NUMBER OF SEQ ID NOS: 14		
	SOFTWARE: PatentIn version 3.2		
	SEQ ID NO 2		
	LENGTH: 793		
	TYPE: PRF		
	ORGANISM: Homo sapiens		
	US-11-060-914-2		
	Alignment Scores:		
	Pred. No.: 4,846-08		Length: 793
	Score: 266.00		Matches: 120
	Percent Similarity: 43.55%		Conservative: 103
	Best Local Similarity: 23.44%		Mismatches: 199
	Query Match: 10.70%		Indels: 90
	Gaps: 7		Gaps: 20


```

Db      53 LeuAaerPhIArGmet-----GluIlethnIleArGAsnSerProTyArGArg 69
QY      181 GAGAGACGCGCTAAAGAAAAGTTGCAGAACACAAACG-----GATTAGACAA 231
Db      70 GluAerSerMeCGluAerPArGluGluAerAerSerLeuAlaIleAerProGlu 89
QY      232 GAGAGACGCGCTAAAGAAAAGTTGCAGAACACAAACGATTAGAACAGATGACTT 291
Db      90 GluMetSerArGluAerGluAerValHisAerG-----LysAerGluAer 104
QY      292 GCTAAAGAAAAGTTGCAGAACACAAACGATTAGAACAGACGCTCTAAAGAA 351
Db      105 ArGluAerGluAerArGluAerHisAerSerHisAerAerGluAerGluAerVal 124
QY      352 AGTTCGCAAGAACACAAACGCGCTTTAGAACAGAGACGCTCTAAAGAAAAGTTGCA 411
Db      125 LysGluAerGluAerG-----GluHisAerGluAerGluAerHisAerGluAer 140
QY      412 GAAACAACAAGCGATTAGAACAGAGACGCTTAAAGAAAAGTTGCAGAACACAA 471
Db      141 GluAerPylAerAerArGluAerGluAerGluAerGluAerAerGluAerHis 160
QY      472 AGCGATTTAGAACAGAGACGCTGCTAAAGAAAAGTTGCAGAACACAAACGATT 531
Db      161 Ser-----ArGluAerGluAerAerGluAerGluAerGluAerGluAer 175
QY      532 GAAACAACAAGCGCTCTAAAGAAAAGTTGCAGAACACAAACGATTAGAACAGAG 591
Db      176 GluAerGluAerGluAerMeArGluGluAerGluAerGluAerGluAerGluAer 195
QY      592 AGACGCTGCTAAAGAAAAGTTGCAGAGACGCAAAAGATTTAGAACAAAGAAAGCTGAT 651
Db      196 ArGluAerGluAerGluAerGluAerGluAerGluAer----- 206
QY      652 ACGAAAAAAATTTAGAAAAAGAAAAGAAATGAGATATATTAGCAGAGATTATAT 711
Db      206 ----- 206
QY      712 GGTCTTTAGAAAATACCAGCTATAGAACTTCATCAGAAAAACGTCGATATTATATA 771
Db      207 AlaAerGluAerValSerA----- 213
QY      772 CCACATCAATCTCTTACCTCAGACAAACAGAGGAATAGTAGATTCCAAGAAATA 831
Db      214 ---HisHisAerGluAerHisAerGluAerPylAerValLysAerHisAerHisAer 232
QY      832 TCTATATATGAAAAACAATAGAAATCTTATCAACAATGTTGAAGACGAGAGGAT 891
Db      233 ArGSerProProArGProProArGluAerGluAerGluAerGluAerGluAer 251
QY      892 ATACATTAAGACATCTTGAAGAAAAGAAAGATGCTCAATAAACCAACAAAGAAAGA 951
Db      252 GluGluAerAerProAerAerGluAerGluAerProAerGluAerGluAerMeArGlu 271
QY      952 GATTAATCT-----GTCGACATACAAATCATACATTAGACAGCACTAATATTCTGAT 1005
Db      272 GluAerAerPylAerAerPylAerGluAer-----Aer 281
QY      1006 GTTAATGATTTTCAATAAGTAAGTATGAGATGAATAAGTCTGAATATGACGATTCA 1065
Db      282 HisAerAerPylAerGluAerHisAerSerHisAerAerGluAerGluAerGluAer 301
QY      1066 TTAATGATGAAGAAAGATGATGAGACTTAGAAGATTTAAGCTTATGCGCAATAT 1125
Db      302 GluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAerHisAerGlu 321
QY      1126 GACAATTTCAAGATGAGAAACATAGCAATTTATTAAGAACTAAGAAAGTTGATAGAG 1185
Db      322 GluAerGluAerGluAerGluAer-----GluAerGluAerGluAer 333
QY      1186 AAAAAATTAATTTAGATTTAGATGAAGAAATCAATCAGAAAGATTAATCT 1245

```

```

Db      334 GluGluAerGluAerGluAerSerAerAerSerGluAerAerGluAerGluAerValSer 353
QY      1246 GAGAAAAATTAATAAGAAAAGAAAGAAATATGAAAAACAAGATTAATTTAAACCA 1305
Db      354 GluGluAerGluAerSerGluAerPylAer-----GluAerGluAerGluAerHisAerVal 371
QY      1306 AATGATTAAGTTGATGATGAGCATATTAATAATTTAAATAATGATAGCAGGTTAAT 1365
Db      372 ValProGluAerAerPylAerPylAer-----AerAerSerGluAer 384
QY      1366 AAGAAAAAGAAAATTCATTAATCATTTCTTATATTAGACGAGACAAAGAAAT 1425
Db      385 GluGluAerGluAer-----GluAerGluAerGluAerHisAerGluAer 398
QY      1426 TTACAGATCGTGAT-----GAGTTA 1446
Db      399 SerAlaAerGluAerGluAerAerPylAerProAerSerProAerAerPylAer 418
QY      1447 TCTGAAGATATTAATTAATTTATTAAGAACTA 1479
Db      419 LysGluAerGluAerProAerPylAerProAerAer 429

RESULT 4
US-10-821-234-1143
/ Sequence 1143, Application US/10821234
/ Publication No. US20050255114A1
/ GENERAL INFORMATION:
/ APPLICANT: Stache-Crain, Birgit
/ APPLICANT: Andarmani, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821,234
/ PRIOR FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ NUMBER OF SEQ ID NOS: 1704
/ SOFTWARE: pc_seq_version 1.0
/ SEQ ID NO 1143
/ LENGTH: 1388
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-821-234-1143

Alignment Scores:
pred. No.: 1,82e-07 Length: 1388
Score: 254.00 Matches: 112
Percent Similarity: 42.75% Conservative: 106
Best Local Similarity: 21.96% Mismatches: 144
Query Match: 10.22% Indels: 148
DB: 6 Gaps: 21

US-09-837-344-41 (1-1482) x US-10-821-234-1143 (1-1388)
QY      1 CAAGAACAAACAAGCGATCTAGAA-----CAAGAGACGCTGCTAAAGAAAAGTTGCA 54
Db      582 GluGluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAer 601
QY      55 GACAAACAAGCGATTTAGAACACATAGACTTGTCT--AAAGAAAAGTTACAAGCGAG 111
Db      602 GluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAer 621
QY      112 CAAAGCGATTTAGAACAGAGAGACTGCTTAAGAAAAGTTGCA----- 156
Db      622 AlaAerGluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAer 641
QY      156 ----- 156
Db      642 ValAerGluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAerGluAer 661
QY      157 -----GACAAACAAGC 168

```


Db 680 ArgGlnbenLeuEngInLeuSerGluIleYsGInGInArgLYs----- 695
Qy 610 TTGCAAGACGACAAAGAGATTAGAACAAAGAAAGCTGATACGAAAAAAATTAGAA 669
Db 695 ----- 695
Qy 670 AGAAAAAGAACATGAGATATATAGACAGAGATTATATGTCGTTAGAAATACA 729
Db 696 -----AspLeuAenGInIleYsIleGluIle--- 703
Qy 730 GCTATAGAACTTCATCAGAAATGAACGCGATATTATATACCAATCAATCTTCTTTA 789
Db 704 -----AsnGInSerGInLeuGInIleYsCySInLeGInAspIleLeu 716
Qy 790 CCTCAGCAACAAGAGGAAATGTAAGAGATTCCAAAGAAATATGTTATATGAAAAACA 849
Db 717 SerIleGInAen----- 720
Qy 850 AATGAGAAATCTATTACAAATAATGTTGAAAGACGAAGGATATACATTAAGACATCTT 909
Db 721 -----HisTyr 722
Qy 910 GAAGAAAAAGATGTTCAATTAATAACGAAACAAAGAAATGAATTCGTGACATA 969
Db 723 GInAspIleYs-----AlaIleYsGInSerIleYsLeuAspValLeu 735
Qy 970 CAAATCATCATATAGACAGATAATATTTCTGATGTTATGATTTTCAATAGTAAG 1039
Db 736 ILeuAsnIleAlaIleAspIleLeuAen-----AspTyrIleGInLeuThr--- 750
Qy 1030 TATGAGCATGAAATTAAGTCTGATATATGACGATTCATTATATGATGAAGAAAGATGAT 1089
Db 751 ---ValGInAenGInArgMetGInIleYsIleAsp-----SerAsp 761
Qy 1090 GAAGACTAGACGAATTAAGCTTATGCAATATGCAATTCCTCAAGATGAAGAA--- 1146
Db 762 GInThrIleAspAenLeuArgIleYsValIleYsLeuThrIleYsMetThrIleAspGInLeu 781
Qy 1147 -----AACATAGAAATTTATTAAGAACTAGAAAGATTTGATAGAGAAA----- 1188
Db 782 GInProValAenLeuAenAlaIleGInGInPheGInIleuAenGInIleArgTyrThrPhe 801
Qy 1189 ---AATGAAATTTAGATGATTTAGATGAGAAATGAAATATCTCGAAAGATTAATCT 1245
Db 802 LeuAenGInGInArgThrAspLeuArgGInIleYsGInIleuGInIleIleHis 821
Qy 1246 GAAGAAAAATTAATAAGAAAGAAATATGAAAGAAAGATTAATTTAAACCA 1305
Db 822 GIn-----MetAspIleYsGInValGInGInIleYsArgPheYsThr 833
Qy 1306 AATGATAAAGTTTGTATGATGAGCATATTAATAAATATTAATAATGATAGCAGGTTAAT 1365
Db 834 ThrPheIleAlaValGInAspIleAspPheThrValPheYs----- 847
Qy 1366 AAGAAAAAGAAAAATTCATTAATATGATTTCTATATATTGACGAGACAAATGAATTT 1425
Db 848 -----GInLeuPheGInGInIleYsIleGInIleGIn 856
Qy 1426 TTACGATCGTGATGAG 1443
Db 857 LeuArgLeuThrGInAsp 862

RESULT 7
US-11-074-176-134
; Sequence 134, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Kleenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding

; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-134

Alignment Scores:
Pred. No.: 2,97e-06 Length: 1189
Score: 230.50 Matches: 144
Percent Similarity: 40.21% Conservative: 129
Best local Similarity: 21.21% Mismatches: 191
Query Match: 9.27% Indels: 215
DB: Gaps: 33

US-09-837-344-41 (1-1482) x US-11-074-176-134 (1-1189)
Qy 1 CAAGACACAAAGCGATCTAGAACAAAGAGAGA-----CGTGCTAAAGAAAG 48
Db 175 GInGInIleAenGInGInIleYsIleYsThrGInAspAenLeuIleArgIleAenAspLeu 194
Qy 49 TTGCAAGACAAACAAAGCGATTTA-----GAACAGATAGACTTGCTAAAGAA 96
Db 195 ValIleAspIleuAenGInSerIleYsArgLeuGInProLeuAenGInGInSerIleuAlaIleGIn 214
Qy 97 ---AAGTTACAGACGACCAAGCGATTTAGAACAAAGAGACTTGCTAA-- 144
Db 215 TyrIleAspPheGInIleYsSerGInIleuAspIleYsLeuIleYsSerIleuAlaPheGInIle 234
Qy 145 GAAGAGTTGCAAGACAAAGCGATCTAGAACAAAG----- 183
Db 235 GInAenIleAenGInGInArgGInAspIleGInIleYsSerAlaAspIleYsAenIleLeu 254
Qy 184 ---AGACGTGCTAAAGAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGAGAGCT 240
Db 255 LeuAlaIleYsLeuAspAspGInValIleYsAspSerGInAlaValIleThrGInIleYsArgAla 274
Qy 241 GCTTAAGAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGTTCCTAA-- 297
Db 275 GInTyrIleYsIleuAenArgAspGInArg---AspHisThrGInAenIleYsLeuIleYsLeu 293
Qy 298 ---GAAGAGTTACAAAGCGACCAAGAGCGATTTA-----GAACAAAGAGAGCGTCT 345
Db 294 SerIleYsAspLeuSerGInLeuAenAlaSerIleuGInMetAlaGInGInSerArgGInPhe 313
Qy 346 AAAGAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAG 405
Db 314 AspAspAlaThrIleYsGInGInIleYsIleYsAen-----GInValIleYsGInIleYsGInAen 331
Qy 406 TTGCAAGACAAACAAAGCGATTTAGAACAAAGAGAGACTTGCTAAAGAAAGTTGCAAGAA 465
Db 332 LeuValGInIleuAenIleYsAlaAspLeuAspGInLeuIleYsIleYsGInIleYsIleYsGInAsp 351
Qy 466 CAACAAAGCGATTTAGAACAAAGAGAGA-----CGTGCTAA 501
Db 352 GInGInAspValIleuIleYsIleGInArgGInIleuThrGInIleuAenGInAspPro 371
Qy 502 GAAAGTTGCAAGAACAAAGCGAT-----TTAGAACAA 537
Db 372 GInIleuAenIleYsIleYsAspIleArgAsnAenTyrMetGInIleuGInAsp 391
Qy 538 GAGAGCGTGTAAAGAAAGTTGCAAGAACAAACAAAGCGATTTAGAACAAAGAGAGAGCT 597
Db 392 GInAlaThrThrAenAenGInIleValAenIleuAenSerAspLeuArgArgSerGInAla 411
Qy 598 -----GCTAAAGAAAGTTGCAAGAGAGAGCAAGAGATTTA 633

```
Db 412 AspThrThrTyrGlnThrGlyAspValSerIysGlnLeuThrAspAlaGlnIlySerGlnLeu 431
Qy 634 GAACAAAGAGAGCGGTGATACGAAAAAATTGAA--AGAAAAAGACATGAGAGAT 690
Db 432 GluGlnLeuArgIleGluGlyLysIlySerLeuThrAspLysArgGlnIlySerGlnAsnAla 451
Qy 691 ATATTA----- 696
Db 452 IleValArgIleAsnIlySerGlnAsnGlnLeuAsnThrGlnLeuThrAsnLeuArgGln 471
Qy 697 -----GCAGAGATTATATGCGCTTGAAGATACAGCATAGACATTCATCA 747
Db 472 ValIleAsnAlaGlu-----ArgAsnGlnLeuGlnLysValGlnAlaArgHis 487
Qy 748 GAA-----AATGAACGTGATATTATATACCATCAATCTTCT 786
Db 488 GluAlaLeuValAsnIleGlnIlySerArgHisGluGlyTyrTyrGlyValArgAsnVal 507
Qy 787 TTA----- 789
Db 508 LeuAsnHisLeuAsnAspPheAlaGlyValIleGlyAlaValGlyGlnLeuIleThrPhe 527
Qy 790 -----CCTCAGAGACAACAGA 804
Db 528 ProAlaGlnLeuGlnAlaAlaMetThrThrAlaLeuGlyGlyValGlnIlePhe 547
Qy 805 GGGAGATAGTAGAGATCCAGAGAAATATCTATATATGAA--AAAACAAT----- 852
Db 548 ThrGluSerArgIleSerAlaArgAsnAlaIleAsnIlySerLeuGlnAsnHisGlyGly 567
Qy 853 -----AGAGATCTATT 864
Db 568 ArgAlaThrPheLeuProLeuAspGlyLeuArgGlnTyrGlyIleProGlnSerThrVal 587
Qy 865 ACACACA-----AATGTTGAAGCAGAGGATATATACATAAAGACATCTTGAAGAAAG 918
Db 588 ThrThrLeuLysSerIlyrAspGlyPheArgGlyIleAlaSerAspLeuValGluSerIys 607
Qy 919 AAAGATGTTCAATTAACCCAGAACAAAAGATTAATCTGCTGACATCAAAATCAT 978
Db 608 ThrAsp-----GlnAspIleThrAlaAlaIle--AsnTyr 618
Qy 979 ACATTAGACAGACATAATTTCTGAT----- 1005
Db 619 LeuLeuGlySerValIleValIleValAspThrIleAspThrAlaMetSerValAlaGlnArg 638
Qy 1006 GTTAATGATTTTCAATTAAGTAAGTATGAGATGAATTAAGTCT----- 1050
Db 639 ValAsnAspGlyrArgIleValIleThrLeuAspGlyAspValIleSerProGlyGlySerMet 658
Qy 1051 -----GATATGACGATTCATTAATGATCAAGAGACATGATGAA 1092
Db 659 ThrGlyGlyGlnArgAsnGlnArgSerAsnSerProLeuGlnThrAlaThrGlnIleAsn 678
Qy 1093 GACTTGAAC--GAATTTAAGCCTATTGCGCATATGACATTTCCAGATGAAGAAAC 1149
Db 679 GlnLeuGlnIlySerGlnIleLysThrLeuLysGln-----AsnLeuAsnGlnIlyAspGlnAsp 696
Qy 1150 ATAGCAATTTATAAAGACATGAAAGATTGATGAGAAA-----AATGAAGAT 1197
Db 697 -----LysLeuGlnIleAsnLeuValAspGlnSerAsnLysValAsnAlaGln 711
Qy 1198 TTAGATGATTTAGATGAAGAAATGAAATTCATCAGAGATTTATCTGAAGAAAAAT 1257
Db 712 LeuGlnAspLeuGlnIleAspAlaLeuArgGlnThrSerGlnAlaIleAsnGlnIleAlaIle 731
Qy 1258 AAA-----AAAGAAAGAAATATGAAGAAAAAAGAGATTAATTTAAACCAAT 1308
Db 732 SerPheGlnIlyGlnIlyGlnIlyValLysArgLeuLeuAspAlaAsn----- 747
Qy 1309 GATAAAGTTTGTATGATGACATATTTAA----- 1338
```

```
Db 748 -----ThrLeuTyrLysSerArgIleLysAspArgAsnAspArgIleGlnLeuLeuLys 765
Qy 1339 -----AAATATTAATAATGATTAAGCAGGTT-----AATAAGAA 1371
Db 766 LysGlnIleLysGlnAlaAsnAspLysGlnMetLeuThrIlyGlnIlyGlnIlyGln 785
Qy 1372 AAGGAAAAATTCATTAATTCATTTGTTGATATATTATGACGAGACAAATGAAATTTACAG 1431
Db 786 LysAlaLysMetAsnAspLeuGlnAspLysIleLysAsnPheAsnAsnLeuSerGlnArg 805
Qy 1432 ATCGGATGATGATTATCTGAA-----GATATTAATAATTTATATGAAACTA 1479
Db 806 IleGlnAspGlnLeuSerLysLeuAspProIlyAlaValIlyThrAsnLysLeu 824

RESULT 8
US-10-821-234-901
; Sequence 901, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 901
; LENGTH: 1586
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-821-234-901

Alignment Scores:
Pred. No.: 3,816-06 Length: 1586
Score: 228.00 Matches: 126
Percent Similarity: 38.30% Conservative: 95
Best Local Similarity: 21.84% Mismatches: 168
Query Match: 9.17% Indels: 168
DB: Gaps: 23

US-09-837-344-41 (1-1482) x US-10-821-234-901 (1-1586)
Qy 1 CAAGAACAAACAAGCGATCTAGACACAGACAGCGTGAAGAAAG-----TTGCA 54
Db 1025 GluGlnGlnArgLysAlaIleGlnIleLysValAlaAlaPheGlnIlySerGlnValIleGln 1044
Qy 55 GAACAAACAAGCGATTTGAACAAGATGACCTTGTAAGAAAGTACAGAG----- 108
Db 1045 LeuGlnAlaSerHisArgIlySerGlnIlyAlaLeuGlnIlyArgLeuAspGlnValSer 1064
Qy 109 -----CAGCAAGCGATTTAGAAACAAGACAGACTTGCTAAAGAAAGTTG 153
Db 1065 ArgGlnLeuCyHisIleThrGlnIleSerHisAlaSerLeuArgAlaAspAlaGlnVal 1084
Qy 154 CAAGAACAAACAAGCGATCTA-----GACAAAG 183
Db 1085 GlnGlnGlnGlnGlnGlnIleMetAlaGlnLeuHisSerLysLeuGlnIleSerGlnValIleGln 1104
Qy 184 AAGCGTGTAA-----GAAGAATTGCAAGAACAAACAAGCGATTTGAACAAGACAG 240
Db 1105 ValArgSerLysCySerGlnIlyLeuSerGlyLeuHisGlyGlnLeuGlnIlyAlaArgAla 1124
Qy 241 GCTAAAGAAAGTTGCAAGACAAACAAGCGATTTGAA-----CAA 282
Db 1125 GluAsnSerGlnLeuThrGlnArgIleArgSerIleGlnAlaLeuLeuGlnIlyGln 1144
Qy 283 GATAGACTTGCTAAAGAAAGTTTACAGAGACAGCAAGCGATTTAGACACAGACAGCGT 342
```

Db 1145 AlaArgbpa1a61nabp---ValGlna1a5erGln1a61n1a6pGlnGln1nThr 1163
Qy 343 GCTAAAGAAAAGTTGCAGAACAAAGCGATTAGAACAGAG-----AGA 390
Db 1164 ArgLeuYb61nLeuGlnSerGlnValSerG1yLeuG1nVala1a1eG1uLeuArg 1183
Qy 391 CGTCTAAAGAAAAGTTGCAGAACAAAGCGATTAGAACAA-----435
Db 1184 Glna1aVala1aG1nGlnGlnVala1y6a6n6n6pLeuArg1u1y6a6nTPrLy6a1a 1203
Qy 436 -----GAGAGACTTGCTTAAGAAAAGTTGCAGAA-----465
Db 1204 MecGlna1aLeu1a1aThr1a1aGlnGln1a1aCyl6y6eG1u1y6eUhi1eSerLeuThrGln 1223
Qy 465 -----465
Db 1224 Ala1y6eG1uGlnSerG1u1y6eGlnLeuCy6eUhi1eGln1aGlnThrMecGlna1aLeu 1243
Qy 466 -----CAACAAAGCGATTAGAACAGAGAGA 492
Db 1244 Leu1a1aLeuLeuProG1uLeuSerValLeu1a1aGlnGln1n6n1yTrhG1uTrpLeuGln 1263
Qy 493 CGTCTAAAGAAAAG-----TTGCAGAACAAACA 522
Db 1264 AspLeuYb61nYb61yProThrLeuLeuYb61nSerProAlaProAlaGlnProSer 1283
Qy 523 AGCGATTAGAACAGAGACGTCTTAAGAAAAGTTGCAGAACAA---CAAGCGAT 579
Db 1284 Ser6pLeu1a1aSer1yLeu6a1aArg1u1aG1nG1nUtrhGlnSerThrLeuGlna1aGln 1303
Qy 580 TTAAACACAGAGACCGTCTTAAGAAAAGTTGCAGAGCAGCAAGAGATTAGAACAA 639
Db 1304 Cy6a6pGln1yTrA1y6eSer1eLeu1aG1nUtrhG1nUtrhG1nUtrhG1nUtrhG1n 1322
Qy 640 AGGAGCGTGTATACGAAAAAATTTAGAAAAGAAAAG-----678
Db 1323 Ly6eSerVala1G1nG1nGlnGlnVala1TrpArg1a1a1y6eVala1a1aG1nG1n 1342
Qy 679 -----GAACATGAGATATATTAGCAAGAGATTATATGCT 714
Db 1343 LeuGlnYb6e1y6eArgVala1ThVal1y6eUhi1eUlnG1nUtrhVala1a1y6eUlnG1y 1362
Qy 715 CGTTAGAAAATACCA-----GCTATAGAACTTCATCGAAAATGAA 756
Db 1363 GlnUlnGlnSerSer6pGlnVala1aArg1u1n1eThrLeuUhi1eUlnGlna1aG1nUln 1382
Qy 756 -----756
Db 1383 Ly6h1eMec1a1a1a1a1aSer1a1aG1nUtrhG1nUtrhG1nUtrhG1nUtrhG1n 1402
Qy 757 CGTGATATATTATACCAATCAATCTTTTACCCTCAGACAAAGAGGAAATAGTAGA 816
Db 1403 ArgGlnLeuLeuLeuGlnSerGlnSerGlnSerGlnSerGlnSerGlnSerGlnSerGln 1422
Qy 817 GATTCCAGAGAAATATCTTAATAGAAAAAACAATAGAGAAATCTATTACAACAATGTT 876
Db 1423 GlnSer6pGlnUlnUlna1aLeuVala1aArg1nGlnLeuSerGlnUln-----1437
Qy 877 GAAGAGCAGAGGATATACATTAAGGACATCTTGAAGAAAAGAAAGATGTTCAATA--- 933
Db 1438 -----Ly6eSerUhi1eGln-----AspG1yAsp11e1a 1447
Qy 934 -----AAACAGAACAAAGAAAGAAAGATTAATCTGCTGACATCAAAATCAT 978
Db 1448 G1y1a1aProAlaSerSerProGlna1aProProAlaGlnUlna1aProVala1nLeuY6 1467
Qy 979 ACATTAGAGACAGTAAATATTTCTGATGTTATGTTTTCAAATAAGTAAGTAGAGAT 1038
Db 1468 ThrGlnLeuGlnUtrhTrhGlna1a1eLeuGlnUlnGlnUlnGlnUlnGlnUlnGln 1485
Qy 1039 GAAATTAAGTGTGAATATGATGATTCATTAATAGATGAAGAAAGATGATGAAGACTTA 1098
Db 1486 Ly6eUlnTrh1a1aG1nUlnGlnUlna1aGlnThrSerAlaCy6a1y6eUlnGlnUln 1505

Qy 1099 GACGAATTTAAGCTTATTTGCAATATGCAATTTCCAGATGAGAAAGAAACATAGGAATT 1158
Db 1506 GlnUlnYb61nYb61nYb61nYb61nYb61nYb61nYb61nYb61nYb61nYb61n 1521
Qy 1159 TATTAAGAACTAGAGATTTGATAGAGAAAATGAAAATTTAGATGAT-----1206
Db 1522 AlaSerGlnLeuYb61nYb61nYb61nYb61nYb61nYb61nYb61nYb61nYb61n 1541
Qy 1207 -----TTAGATGAGAAATAGAAAATTCATCGAAGAAATTTATCTGAAGAAA 1254
Db 1542 Ala1a1aThrArgLeuGlnGlnGlnUlnLeuYb61nYb61nYb61nYb61nYb61n 1561
Qy 1255 -----ATAAAAA---GGAAGAAATATGAAAAAACAAGGATATATAT 1296
Db 1562 AspThrVala1y6eUlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1578

RESULT 9
US-10-878-556A-136
; Sequence 136, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 136
; LENGTH: 1410
; TYPE: PRF
; ORGANISM: Homo sapiens
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/trbl_human
; DATABASE ENTRY DATE: 2003-02-28
US-10-878-556A-136

Alignment Scores:
Pred. No.: 6,99e-06 Length: 1410
Score: 223.00 Matches: 122
Percent Similarity: 39.13% Conservative: 103
Best local Similarity: 21.22% Mismatches: 186
Query Match: 8.97% Indels: 164
DB: Gaps: 23

US-09-837-344-41 (1-1482) x US-10-878-556A-136 (1-1410)
Qy 1 CAAGAACAAAGCGATTAGAACAGAGACGTGCTTAAGAAAAG-----TTGCAA 54
Db 849 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 868
Qy 55 GAACAAAGCGATTAGAACAGATGAGACTTGAAGAAAAGTTCAAGAG-----108
Db 869 LeuGln1aSerUhi1aArg1uSerGlnUlna1aLeuGlnUlnYb61nYb61nYb61n 888
Qy 109 -----CAGCAAGCGATTGAAACAAAGAGACTTGAAGAAAAGTTG 153
Db 889 ArgGlnUlnCy61nYb61nYb61nYb61nYb61nYb61nYb61nYb61nYb61nYb61n 908
Qy 154 CAAGAACAAAGCGATCTA-----GAACAGAG 183
Db 909 GlnGlnUlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 928
Qy 184 AGAGCTGCTAAA-----GAAAAGTTGCAAGAACAAAGC 219
Db 929 Vala1aArgSer1y6eYb61nGlnUlnLeuSerG1yLeuUhi1eG1nUlnGlnUlna1aArg1a 948
Qy 220 GATTAGAACAA-----GAG 234
Db 949 GlnUlnSerGlnUlnUtrhGlnUlnArg1e1aArgSer1eGlna1aLeuLeuGlna1aG1n 968
Qy 235 AGAGCTCTAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGTATGACTTGCT 294

[illegible]

Oy	1288	GATATAATTATTGAACCAATG-----ATAAACTTGTAATGAGC	1323
Dd	890	Aaplyelyrs-IleasnIumetSerlysgluValThrLyseLubedlualeuense	909
Oy	1330	CATA-----TTAAATATATAAAT	1350
Dd	909	rIueSeerInLeuserYrserThrsrSerSerlysrArglnSerclngIneuglual	929
Oy	1351	GATNAGCAGGTTCATTAAGAAGAAAAGAAAAATTCATAAATCATTTGTTTCATATATTTGCAC	1410
Dd	929	aLeuGlngIngluValysgInleuGlnaengInleuaIagIucelylsyglnHnscI	949
Oy	1411	GGAGACAATGAAATTTTACAGATCGTGGAGATTA	1447
Dd	949	ngluvalIleservallYrrarWethIsleuLeuyr	961
 RESULT 11 US-10-821-234-1050 ; Sequence 1050, Application US/10821234 ; Publication No. US200502551141 ; GENERAL INFORMATION: APPLICANT: Labat, Ivan APPLICANT: Stache-Crain, Birgit APPLICANT: Andarmani, Susan APPLICANT: Tang, Y. Tom TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia FILE REFERENCE: 821A CURRENT APPLICATION NUMBER: US/10/821,234 CURRENT FILING DATE: 2004-04-07 PRIOR APPLICATION NUMBER: US 60/462,047 PRIOR FILING DATE: 2003-04-07 NUMBER OF SEQ ID NOS: 1704 SOFTWARE: pc_seq_genes Version 1.0 SEQ ID NO 1050 LENGTH: 1410 TYPE: PRT ORGANISM: Homo sapiens US-10-821-234-1050			
 Alignment Scores: Pred. No.: 1,69e-05 Length: 1410 Score: 215.50 Matches: 85 Percent Similarity: 48.82% Conservative: 60 Best Local Similarity: 28.62% Mismatches: 107 Query Match: 8.67% Indels: 45 DB: 6 Gaps: 14			
US-09-837-344-41 (1-1482) x US-10-821-234-1050 (1-1410)			
Oy	1	CAAGAACACAAACCGCTTTCAGAACAGACAGACGCTCTTAAGAAAGTTCGAAGACAA	60
Dd	1086	GluuysgIuuysAlaaYgleuGlugluValgtluArghybargrg----	Glugin 1103
Oy	61	CAAGCGCATTTAGAACAAAGATAGA-----	84
Dd	1104	GIuArghysleuGIylGImeArgrThrlnSerIeusnProAlaPrpheserProeu	1123
Oy	85	CTTGCTTAAGAAAAGTTACCAAGACGACCAAGCGATTAGAA-----CAAGAGACCTT	138
Dd	1124	ThrlaGIngmectYprroGluTyrsrberThrluengInArgrProGInguThrVal	1143
Oy	139	GCTAAAGAAAAGTTCGAAGAACCAACAAAGCGATCTTAACCAAGAGAGACGTGCTTAAGAA	198
Dd	1144	IleArghU--LeuGInPproGInglInPArArgrThrllegluArghArpleuGln	1162
Oy	199	AAGTTGCAAGAACCAAGCGATTTCAGAACAA-----	231
Dd	1163	TyrIleThrValserLyrgluGluLeuserSerclYThrValCysProProAprProTrp	1182
Oy	232	GAGAGACCTGCTTAAGAAAAGTTCGAAGAACCAACAAAGCGATTTCAGAACAGATAGCTT	291
Dd	1183	LysArgrApalArAlayrgIuuysleuGluValyeglInglInImechIslEvaIAspmecLeu	1202

```
QY 292 GCTAAGAAAGTTACAAGACGACGATTTAGAACAGAGACGTCCTAAGAA 351
   |||||
Db 1203 SerIysGlu---IleGlnIleuGlnSerIysProAspArgSerIaGluGlnSerAsp 1221
   |||||
QY 352 AAGTTCAGAAACAACAAGCGATTTAGAACAGAGACGTCCTAAGAAAGTTGCA 411
   |||||
Db 1222 ArgIeuArgIysIeuMetIeuGlnIleuGlnPheGlnIleuGlnIleuGlnSerIysGln 1241
   |||||
QY 412 GAACAACAAGCGATTTAGAACAGAGACGTCCTAAGAAAGTTGCAACAACA 471
   |||||
Db 1242 LysAspGluAspArgIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 1261
   |||||
QY 472 AGCGATTTAGAACAGAGACGTCCTAAGAAAGTTGCAAGAA---CAACAAGCGAT 528
   |||||
Db 1262 Arg---LeuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 1280
   |||||
QY 529 TTA-----GAACAAGAGACGTCCTAAGAAAGTTGCAAGAAACAACAAGCGATTTA 582
   |||||
Db 1281 LeuIeuGlnAspGluIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 1299
   |||||
QY 583 GAACAAGAGACGTCCTAAGAAAGTTGCAAGAGACGACAAAGAGATTTAGAACAAAG 642
   |||||
Db 1300 GluIleuGlnAspArgIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 1318
   |||||
QY 643 AAGCGTCATACGAAAGAAATTTAGAAAGAAAGAAAGCATGAGATATATTAGACAGAG 702
   |||||
Db 1319 -----ThrLysArgSerAspIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 1331
   |||||
QY 703 GATTATATGTCGCTTTAGAAATATACCAAGCTATAGAACTTCATCAGAAAGTGAAGCTGGA 762
   |||||
Db 1332 GlyIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 1351
   |||||
QY 763 TATTATATCCACAT-----CAATCTTCTTACCTCAGAC 798
   |||||
Db 1352 LeuIeuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 1368
   |||||

RESULT 12
US-10-467-657-694
/ Sequence 694, Application US/10467657
/ Publication No. US20050260581A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MIGNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/467,657
/ PRIOR FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqMin99, version 1.04
/ SEQ ID NO 694
/ LENGTH: 558
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657-694

Alignment Scores:
Pred. No.: 3 52e-05 Length: 558
Score: 210.50 Matches: 128
Percent Similarity: 41.12% Conservative: 99
Best Local Similarity: 23.19% Mismatches: 188
Query Match: 8.47% Indels: 137
DB: 6 Gaps: 26

US-09-837-344-41 (1-1482) x US-10-467-657-694 (1-558)
QY 1 CAAGAACAACAAGCGATCTAGAACAGAGACGTCCTAAGAAAG----- 48
   |||||
   |||||
```

```
Db 33 GlnAspGluAsnIleuValIleuSerAsnArgPheSerIleuValIleuGlnIleuValIleu 52
   |||||
QY 49 TTGCAAGAAACAACAAGCGATTTAGAACAGATGCTTGAAGAAAGTTACAAGAG 108
   |||||
Db 53 LeuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 72
   |||||
QY 109 CAGCAAGCGATTTA-----GAACAAGAGACGTCCTAAGAAAGTTGCAACAACA 162
   |||||
Db 73 LysIleuGlnIleuValIleuValIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 92
   |||||
QY 163 CAAGCGATCTAGAACAGAGACGTCCTAAGAAAGTTGCAAGAAACAACAAGCGAT 222
   |||||
Db 93 LysIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 112
   |||||
QY 223 TTAGAACAGAGACGTCCTAAGAAAGTTGCAAGAAACAACAAGCGATTTAGAACAA 282
   |||||
Db 113 LeuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 132
   |||||
QY 283 GATGACCTTGTAAAGAAAG-----TTACAAGAGACGACAAAGCGATTTAGAA 330
   |||||
Db 133 ArgPheIleuValIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 152
   |||||
QY 331 CAAGAGACGTCCTAAGAAAGTTGCAAGAGACGACAAAGCGATTTA-----GAACA 384
   |||||
Db 153 ArgIeuArgIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 172
   |||||
QY 385 GAGAGACGTCCTAAGAAAGTTGCAAGAAACAACAAGCGATTTAGAACAGAGACGCT 444
   |||||
Db 173 GluIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 184
   |||||
QY 445 GCTAAGAAAGTTGCAAGAAACAACAAGCGATTTAGAACAGAGACGTCCTAAGAA 504
   |||||
Db 185 AlaSerGlu-----GluIleuSerIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 198
   |||||
QY 505 AAGTTCAGAAACAACAAGCGATTTA-----GAACAAGAGACGCT 546
   |||||
Db 199 SerIeuSerAspGlnPheGlnAsnIleuValIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 218
   |||||
QY 547 GCTAAGAAAGTTGCAAGAAACA-----CAAGCGATTTAGAACAGAGAGAG 594
   |||||
Db 219 PheThrGlnIleuGlnAsnArgIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 238
   |||||
QY 595 CGTGTAAAGAAAGTTGCAAGAGACGACAAAGAGATTTAGAACAAAGAGCGTCGATACG 654
   |||||
Db 239 HisGlnPheGlnIleuValIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 258
   |||||
QY 655 AAAAAAATTTAGAAAGAAAAAGAA-----CATGCAATATATTATTA 696
   |||||
Db 259 LeuGlnAsnGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 278
   |||||
QY 697 GCA-----GAGCATTTATATGTCGCTTTA 720
   |||||
Db 279 AlaIeuThrAsnAlaIeuThrGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 298
   |||||
QY 721 GAATATACAGCTATAGAACTTCATCAGAAATGGAACGTTGA---TATTATATACCAAT 777
   |||||
Db 299 IleuGlnIleuValIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 318
   |||||
QY 778 CATCTTCTTTACCTCAGAGACAGAGGGAATAGTAGA----- 816
   |||||
Db 319 AlaAlaSerValIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 338
   |||||
QY 817 -----GATTCAGAGAAATATCTATATAGAAACAAATAGAGATCTATT 864
   |||||
Db 339 ValAsnIleuProAspAsnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 355
   |||||
QY 865 ACNACAAATGTTGAAGAGCAGAGGATATATACAAAGCATCTTGACAGAAAGAAAGAT 924
   |||||
Db 356 ThrIleuValIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 375
   |||||
QY 925 GGTCAATTAACCAAGACAAAGAAAGAGATTAATGCTGCAATACAAATCATATACA--- 981
   |||||
Db 376 AlaIleuIleuSerIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleuGln 395
   |||||
```

QY 982 TTAGACACAGTAAATATTCTGATGTT----- 1008
DB 396 LeuGlulGlyValLeuThrLeuSerPheValPheMetPheIleProValGluProAlaTyr 415
QY 1009 -----AATGAT-----TTTCAATTAAGTAAAGTATGAGATGAATA 1044
DB 416 LeuLeuAlaLeuGlnAenMetPheAlaGlyLeuPheGln---GluCysPheAspLysArgIle 434
QY 1045 AGTCCTGAATATGACGATTCATTAATA----- 1071
DB 435 MetLeuValGlyProSerThrLeuLeuAlaThrLeuArgThrValAlaAenIleTyrArg 454
QY 1072 GATGAAAGACAGACGTGATGAAGACTTA-----GACCAATTAAGCCATATGTGCA 1122
DB 455 AenGluGlnGlnAenGlnAenAlaLeuAlaIleAlaAspGluGlyLeu----- 472
QY 1123 TATGACAAATTTCCAAAGATGAAGAAACATAGAAATTTATAAAGAACTAGAAAGATTGATA 1182
DB 473 TyrAspLysPhe-----ValGlyPheValGlnThrLeuGluSerVal--- 486
QY 1183 GAGAAATATGAAATTTAGATGATTTAGATGACAAATAGAAATGATCAGAAATTA 1242
DB 487 -----GlyLysGlyIleAspGlnAlaGlnAsnSerPheGlnThrAlaPheLysGlnLeu 504
QY 1243 TCTGCA-----GAAAA-----ATAAA 1280
DB 505 AlaGluGlyArgGlyAenLeuValGlyArgAlaGluLysLeuArgLeuGlyValLys 524
QY 1261 AAAGGAAAGAAATAT-----GAAAAACAAAGATAT 1293
DB 525 AlaGlyLysGlnLeuGlnAenArgLeuValGluArgAlaAenGluGlnThrAlaPheAla 544
QY 1294 AATTTAAACCAATGATTAAGTTTATGATGAG 1329
DB 545 LeuGlyLysProSerGluGlnGluAlaAlaAspGlu 556

RESULT 13
US-10-821-234-1145
; Sequence 1145, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labtec, Ivan
; APPLICANT: Steache-Crain, Birgit
; APPLICANT: Andarmanl, Susan
; APPLICANT: Targ, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PC SEQ_genes Version 1.0
; SEQ ID NO 1145
; LENGTH: 1299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1145

Alignment Scores:
Pred. No.: 3,27e-05 Length: 1299
Score: 210.00 Matches: 124
Percent Similarity: 35.74% Conservative: 119
Best Local Similarity: 18.24% Mismatches: 217
Query Match: 8.45% Indels: 220
DB: Gaps: 26

US-09-837-344-41 (1-1482) x US-10-821-234-1145 (1-1299)

QY 4 GAACACAAAGCGATCTAGACAAAGAGAGACGTGCTAAAGAAAGTTGCAGAAACAACA 63
DB 340 GluLeuGlnLysArgIleAlaGluMetGluThrGlnLysGluLysIleHisGluAspThr 359

QY 64 AGCGATTTA---GAACAAGATAGACTTCTAAAGAAAGTTACAGAGCAGCAAGCGAT 120
DB 360 LysGluIleAsnGlnLysSerAenIleLeuSerAenGluMetLysAlaLysAenLysAsp 379
QY 121 TTAGAA-----CAAGAGACCTTGCTTAA-----GAAAGTTGCAGAA 159
DB 380 ValLysAspThrGluLysLysLeuAenLysIleThrLysPheIleGluGluAenLysGlu 399
QY 160 CAA-----CAAGCGATCTAGAAACAAGAGACGTCTAAAGAAAGTTGCAGAAACA 213
DB 400 LysPheThrGlnLeuAspLeuGlu---AspValGlnValArgGluLysLeuLysHisAla 418
QY 214 CAAGCGATTTA-----GAACAAGAGACGTGCTAAAGAAAGTTGCAGAAACA 261
DB 419 ThrSerValAlaLysLysLeuGluGlnLysGlnLeuGlnLysAspLysGluValGluGln 438
QY 262 CAACAAGCGATTTAGAACAGATAGACTTCTAAAGAAAGTTACAAAGCAGCAAGC 321
DB 439 PheLysSerIleProAlaLysSerAenAenIleLeuGlnLysThrThrArgAenAen 458
QY 322 GATTTAGACAAAGAGACGTGCTAAAGAA---AAGTTGCAGAAACAACAAGCGATTTA 378
DB 459 AlaLeuGluLysGluLysGluLysGluGluLysLeuLysLeuLysValMetAspSerLeu 478
QY 379 GAACAAGAGACGTGCTAAAGAAAGTTGCAGAAACAACAAGCGATTTAGAAACAAGAG 438
DB 479 LysGlnGlnThrGlnGly-----LeuGlnLysGlnLysGlnLysSerArgGluLysGlu 495
QY 439 AGACTT-----GCTAAAGAAAGTTGCAGAAACAACAAGC 474
DB 496 LeuMetCylPheSerLysSerValAenGluAlaArgSerLysMetLysPheValAlaGlnSer 515
QY 475 GATTTAGAA-----CAGAGACGCTGCTTAA 501
DB 516 GluLeuAspThrLeuLysSerArgHisAenThrAlaValSerGlnLeuThrLysAlaLys 535
QY 502 GAAAGTTGCAGAAACAACAAGCGATTTAGAAACAAGACGTGCT----- 549
DB 536 GluAlaLeuIleAlaLysSerGluThrLeuLysGlnLysArgValAlaIleAspAspIle 555
QY 550 AAAGAAAGTTGCAGAAACAACAAGCGATTTAGAAACAAGACGTGCTTAAAGAAAG 609
DB 556 GluLysLysLeuProGlnThrGlnGlnGlnLysGlnLysGlnLysGlnLysGlnLys 575
QY 610 TTGCAGAGACGAA-----AGAGATTAGAAACAAGAGAGCT 648
DB 576 LeuThrGlnGlnGluThrAenPheLysSerLeuValHisAspLeuPheGlnLysValGlu 595
QY 649 GATTCGAAAAAATTTAGAAAGAAAAAAGAAACATGACATATATTA----- 696
DB 596 GluAlaLysSerSerLeuAlaMetLysAsnArgSerArgLysValLysLeuAspAlaIleIle 615
QY 697 -----GCAGAGATTATATGCTGCTTA----- 720
DB 616 GlnGluLysLysSerGlyArgIleProGlyIleTyrGlyArgLeuGlnLysAspLeuGlyAla 635
QY 720 ----- 720
DB 636 IleAspGluLysTyrAspValAlaIleSerSerCysHisAlaLeuAspTyrIleVal 655
QY 720 ----- 720
DB 656 ValAspSerIleAspIleAlaGlnGluCysValAsnPheLeuLysArgGlnAenIleGly 675
QY 721 -----GAAATACCGACTATA 735
DB 676 ValAlaThrPheIleGlyLeuAspLysMetAlaValThrAlaLysLysMetThrGluIle 695
QY 736 GAATTCATCATGAAAAATGAAGCT----- 759
DB 696 GlnThrProGluAenThrProArgLeuPheAspLeuValLysValLysAspGluLysIle 715

```

QY 760 -----GGATATTATACACATCATCTCTTTACTCGACAGACAGAGGAAATAGT 813
Db 726 ArgGlnAlaPheTyrPheAlaLeuArgPheThrLeuValAlaAspAsnLeuSerGlnAla 735
QY 814 -----AGAGATTCGAAAG-----GAAATATCTATA 837
Db 726 ThrArgValAlaTyrGlnLysAspArgTyrPArgValAlaThrLeuGlnGlnLysGlnIle 755
QY 838 AATGAAAAAACAATATAGAAATCTATTACAAATGTTGAAAGGACGAAGGATATACAT 897
Db 756 IIEGTLGINSERGILY-----ThrmethrGlyGlySerLysValMet 770
QY 898 AAGAGACATCTTGAAGAAAAAGAGTGTTCATTAATAACCAAAACAAAGAAAGATAA 957
Db 771 LysGlyArgMetCylSerSerLeuValIIEGTLIESerGlnGlnLysValAsnLysMet 790
QY 958 TCTGCTGACATACAAAATCATACATTAGACAGACATTAATTTCTGAGTTAATGATTTT 1017
Db 791 GluSerGlnLeuGlnAsnAspSerLysLysAlaMetGlnIle----- 804
QY 1018 CAAATATAGTAATATGAGATGAATAATAGTCTGAATATAGAGATTCAATATA----- 1071
Db 805 -----GlnGlnGlnLysValGlnLeuGlnGlnArgValAlaLysLeu 818
QY 1072 -----GATGAAGAAAGAGATGATGAGACTTAGACGAATTTAAGCTTATGCAAA-- 1122
Db 819 ArgHisSerGlnArgLysMetArgAsnThrLeuGlnLysPheThrAlaSerIleGlnArg 838
QY 1123 -----TATGACAAATTTCCAAAGT----- 1140
Db 839 LeuIIEGTLGlnGlnLysLeuAsnValGlnValLysLeuGlnLysValLeu 858
QY 1141 -----GAGAAAAACATAGCAATTTATTA 1164
Db 859 AlaThrAlaProAspLysLysGlnLysLeuLeuGlnGlnAsnValSerAlaPheLys 878
QY 1165 -----GAACACTGAGATTTGATAGAGAA----- 1188
Db 879 ThrGlnLysArgAlaValAlaGlnLysAlaGlyLysValGlnAlaGlnLysArgLeu 898
QY 1189 -----AATGAAAAATTTAGATGATTAGATGATGAGCAATAGCA 1224
Db 899 HisAsnThrIIEValGlnLysLeuAsnHisLysLeuLysAlaGlnGlnAsnLysLeuAsp 918
QY 1225 AATATCATCGAAGAAATATATCTGAA-----GAAAAATTAATAAGGAAGAAATATGAA 1278
Db 919 LysIIEAsnLysGlnLeuAspGlnLysAlaSerAlaIIEThrLysAlaGlnValAlaIIE 938
QY 1279 AAAACCAAGGATTAATTTTAAACCAATGATTAAGTTGTATGATGAGCATATTA 1338
Db 939 LysThrAlaAspArgAsnLeu-----GlnLysAlaGlnAspSerValLeuArg 954
QY 1339 AATATTAATAATGATTAAGCAAGTTAATAAGAAAAAGAAAAATTCATAAATCATTTGTT 1398
Db 955 ThrGlnLysGlnLysAspThrGlnLysGlnLysAlaAspAspLeuThrAlaGlnLysLeu 974
QY 1399 CATATATTTGACGAGACATGAAATTTTACAGATGCTGATGATGATTAATGATTA 1458
Db 975 SerLeuGlnAspLysAlaAlaGlnLysValIIEAsnThrAsnAlaAlaGlnGlnSerLeu 994

```

```

/ APPLICANT: YANG, Junning
/ TITLE OF INVENTION: CELL SIGNALING PROTEINS
/ FILE REFERENCE: PF-0521 PCT
/ CURRENT APPLICATION NUMBER: US/11/099,691
/ CURRENT FILING DATE: 2005-04-06
/ PRIOR APPLICATION NUMBER: US/09/700,444
/ PRIOR FILING DATE: 2002-08-26
/ PRIOR APPLICATION NUMBER: 60/085,343
/ PRIOR FILING DATE: 1998-05-13
/ PRIOR APPLICATION NUMBER: 60/098,010
/ PRIOR FILING DATE: 1998-08-26
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PERL Program
/ SEQ ID NO 7
/ LENGTH: 898
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc-feature
/ OTHER INFORMATION: Incyte Clone 2948818
US-11-099-691-7

Alignment Scores:
Pred. No.: 4,13e-05 Length: 898
Score: 208.50 Matches: 105
Percent Similarity: 45.34% Conservative: 119
Best Local Similarity: 21.26% Mismatches: 159
Query Match: 8,39% Indels: 111
DB: 7 Gaps: 22

US-09-837-344-41 (1-1482) x US-11-099-691-7 (1-898)
QY 4 GAACAACAAGCGATCTGACAACAAG-----AGACGTGCTTAAGAAAGTTG 51
Db 452 GlnGlnGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 471
QY 52 CAAGAACA-----CAACGATTTAGAACCAAGATGACTT 87
Db 472 GlnLysGlnLysIIEAlaLeuGlnLysLysLysLysLysLysLysLysLysLysLysLys 491
QY 88 GCTAAAGAAAGTTTCAAGAG-----CAGCAACGCGATTTGAACAAGAGAGACTT 138
Db 492 LysLeuGlnLysGlnLysGlnLysThrHisAlaAsnAsnSerIIEGlnLysLeuLysLeu 511
QY 139 GCTAAAGAAAGTTGCAAGAACCAACGCGATTTAGAACCAAGAGAGAGACTT----- 189
Db 512 AlaLys-----LysGlnValAlaIIEIIEGlnLysGlnLysValAlaIIE 527
QY 190 GCTAAAGAAAGTTGCAAGAACCAACGCGATTTAGAACCAAGAGAGAGAGAGAGAG 249
Db 528 AlaAspGlnLysLysPheGlnGlnGlnIIELeuAlaGlnGlnLysLysAspLeuThrThr 547
QY 250 AAGTTGCAAGAACCAACGCGATTTAGAACCAAGTTGCTTAAGAAAGTTTCA 309
Db 548 PheLeuGlnSerGlnLys-----LysGlnLysIIECylLysGlnLysLysIIE 564
QY 310 GAGCAACAAGCGATTTGACAAGAGAGAGCGTGAAGAAAGTTGCAAGAACCA 369
Db 565 GlnGlnLysMetAsn-----GlnAspHisSerThrProLysLysGlnLysGlnLysGln 582
QY 370 AGCGATTTAGAACCAAGAGAGAGCGTGAAGAAAGTTGCAAGAACCAACGCGATTTA 429
Db 583 Ser-----LysHisLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGln 596
QY 430 GAACAAGAGAGACTTGTCAAGAAAGTTGCAAGAACCAACGCGATTTAGAACCAAG 489
Db 597 GlnAlaHisLeuLeuLeuThrGlnGlnLysGlnLysLysLysLysLysLysLysLysLys 612
QY 490 AGACGTGCTTAAGAAAGTTGCAAGAACCAACGCGATTTGAACAAGAGAGAGAGAG 549
Db 613 ArgPheLysArgLysLysIIEMetLIEuLysArgHisGlnLysGlnGlnGlnIIE--AsnIIE 631
QY 550 AAGAAAGTTGCAAGAACCAACGCGATTTAGAACCAAGAGAGAGAGAGAGAGAGAG 609

```

```

Dh      632  ArgGluGluLeuAenLylVsArgThGlnLysGluMetGlnHlSalMetLeuIlleArg 651
Qy      610  TTGCAAGACGACGAAAGAGATTTTGAACAAAGSAGGCTGATACGAAAAAAATTTAGAA 669
Dh      652  HlVaRgRluseRThrArGsluLeuGlnLylTrArGslnLeuHlSethLysGlnLylSleu 670
Qy      670  AGAAAAAGCAATGAGATATTTTATGACAGAGATTTATATGCTGCTTTAGAAATACCA 729
Dh      671  -----ArgMetAspLeuIlle 675
Qy      730  GCTATAGAACTTCATCAGAAATGAAAGCTGATATTATATACCACTCATCTCTTTTA 789
Dh      676  ArgLeuGln-----HlSglnThGlnLeu 683
Qy      790  CCTCAGGACAAACAGAGGAAATAGTAGAGATTTCCAGGAAATATCTATATAGAAAAACA 849
Dh      684  GlnuBnGlnLeuGlnLylTrAenLylArGsluArgGlnuArgGlnuLeu----- 697
Qy      850  AATAGAGAACTTATTACAAACAATGTTGAAAGCAAGGATATATACATTAAGCAATCTT 909
Dh      698  HlVaRgRluseRlValMetGlnLeuArgGlnGlnInProLysAenLeuLylValMetGluMet 717
Qy      910  GAAGAAAAAGAA-----GATGCTCAATTAACACAGAACAAAGAAAGAAAGATTAATCTGCT 963
Dh      718  GlnLlLeuLylVsGlnRheGlnAerThrcLysLylValGlnThRylVsGlnLylLylVal 736
Qy      964  GACATTCACAAATTCATACATTTAGACACAGTAATATTTTCTGATGTTATATGATTTTCAA 1020
Dh      737  --LeuLylVsAenHlSglnLeuGln-----ValThRProLylVsAenGlnLlLeuLylThR 752
Qy      1021  --ATAAGTAAGTATGAGATGAA-----ATAAGCTGTAATATGAC 1059
Dh      753  lLeuLylVsThRLeuLylVsAerGlnGlnInThRArgLylSleuAlleuAlleuAlleuGlnInThR 772
Qy      1060  GATTCATTATATAGATGAAAGAAAGATGATGAACATTTAGACGAATTTTAAGCCATTTGTG 1119
Dh      773  GlnGlnSerlLeuAenGlnLeuMetMetAlaSerGlnAlaLeu----- 785
Qy      1120  CAATATGACAAATTTCCAAATGATGAAGAAACATTAAGAAATTTTAAAGAACTAGAGATTTG 1179
Dh      786  ArgLeuAerGsluAlleuGlnGlnAlleuGlnLylSglnAlleuAlleuArgLeuGlnLeuGln 803
Qy      1180  ATAGAGAAAAATGAAATTTTATGATGATTTAGATGAAGAAATGAAAAAAATCATCGAAGAA 1239
Dh      804  --GlnGlnMetGlnLeuLeuAenAlleuLylTrGlnSerLylSleuLylMetGlnInThGlnAla 822
Qy      1240  TTATCTGAAGAAAAATTTAAAAAGAAAGAAATTTGAAAAAACAAAGATTAATATTTT 1299
Dh      823  GlnHlSglnLylVsGlnLeuGlnLylVsLeuGlnIn----- 833
Qy      1300  AAACCAATATGATTAAGTATTTATGATGATGAGCATTTTAAAAATATTAATAATGATAG-- 1356
Dh      834  -----ArgValSerLeuAerArgAlaHlSleuGlnGlnLylSleuGlnGlnLeu 850
Qy      1357  --CAGGTTAATAGAGAAAGAAAAATTCATTAATAATCATTTG 1395
Dh      851  AlAlaLeuGlnLylVsGlnLylVsSerGlnuArgLlLeuLylVsAenLeu 864

RESULT 15
US-11-085-775-2
; Sequence 2, Application US/11085775
; Publication No. US20050260634A1
GENERAL INFORMATION:
APPLICANT: BALDWIN, DARYL
APPLICANT: CLARK, HILARY
APPLICANT: JUBB, ADRIAN
APPLICANT: KOEPPEN, HARTMUT
APPLICANT: QUAN, CLIFFORD
APPLICANT: WU, THOMAS
APPLICANT: ZHANG, ZEMIN
TITLE OF INVENTION: ACID-DE-SCURE LIKE-2 POLYPEPTIDES AND ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR

```

[illegible]

Mon Dec 12 08:09:37 2005

us-09-837-344-41.rapbn

Page 16

Search completed: December 7, 2005, 07:16:34
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:30:46 ; Search time 5601.15 Seconds
(without alignments)
2639.584 Million cell updates/sec

Title: US-09-837-344-31
Perfect score: 1563
Sequence: 1 SDLEQERRAKKQLQEQSD.....GHLEKKDSIKPEQKEDKS 316

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-O/cgnt2_1/USPFO.spool h/US09837344/runat.05122005.133532.750/app.query.fasta_1.718
-DB=EST -QFMT=fasta -SUPFIX=ret -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09837344 @CGN 1.1 6731 @runat.05122005.133532.750 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUDRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: EST:
1: gb_est1:.*
2: gb_est2:.*
3: gb_est3:.*
4: gb_est4:.*
5: gb_est5:.*
6: gb_est6:.*
7: gb_est7:.*
8: gb_est8:.*
9: gb_est9:.*
10: gb_est10:.*
11: gb_est11:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	25.3	317	10 T09652	T09652 0224m3 gmbp
2	390.5	25.0	777	10 BX200721	BX200721 Danilo rer
3	350.5	22.4	619	8 CX031474	CX031474 1343594 N
4	345.5	22.1	848	9 A2672705	A2672705 ENTMB16TF
5	328.5	21.0	743	10 C2247380	C2247380 ATRA-aaf3
6	325	20.8	718	9 A2972907	A2972907 2M0246M24
7	322.5	20.6	878	9 BH149306	BH149306 ENTPH89TF

8	320.5	20.5	1261	2 BG533805	BG533805 602562478
9	320	20.5	894	9 A2688399	A2688399 ENTFF244TF
10	319.5	20.4	868	9 A256100	A256100 ENTTCO15TF
11	319	20.4	771	7 CN036384	CN036384 nm.15.c6
12	317.5	20.3	650	9 A2337339	A2337339 1M0068B03
13	316.5	20.2	1041	10 CW923205	CW923205 EDCAX27TR
14	316	20.2	612	7 CN508745	CN508745 AGENCOURT
15	314	20.1	743	3 BJ395948	BJ395948 BJ395948
16	308.5	19.7	814	9 A2672545	A2672545 ENTUK44TF
17	307.5	19.7	1327	10 AG429715	AG429715 Mus muscu
18	305	19.5	1264	2 BG541846	BG541846 602565739
19	304.5	19.4	1723	10 AG330464	AG330464 Mus muscu
20	304	19.4	871	9 A2530343	A2530343 ENTWC91TF
21	303	19.4	704	7 CF864218	CF864218 ps25009xi
22	303	19.4	1794	7 AG382218	AG382218 Mus muscu
23	301.5	19.3	604	3 B0264044	B0264044 faa21e04.
24	300	19.2	1205	10 AG278065	AG278065 Mus muscu
25	299	19.1	890	9 BH163382	BH163382 ENTQX21TF
26	299	19.1	1465	10 AG386901	AG386901 Mus muscu
27	297.5	19.0	737	1 AJ724363	AJ724363 AJ724363
28	297.5	19.0	1921	10 AG365436	AG365436 Mus muscu
29	297	19.0	693	7 CK711667	CK711667 ZP201-P00
30	297	19.0	1945	10 AG370924	AG370924 Mus muscu
31	296.5	19.0	1567	10 AG349825	AG349825 Mus muscu
32	295	18.9	670	1 AL725110	AL725110 AL725110
33	295	18.9	1439	10 CG743836	CG743836 P036-1-H0
34	294.5	18.8	1957	10 AG382393	AG382393 Mus muscu
35	294	18.8	1933	4 AY325173	AY325173 Rattus no
36	293.5	18.8	548	11 DR2M2T	AL735260 Danilo rer
37	293	18.7	1409	10 AG390910	AG390910 Mus muscu
38	293	18.7	1969	10 AG330667	AG330667 Mus muscu
39	290	18.6	1425	10 AG310594	AG310594 Mus muscu
40	289.5	18.5	869	7 CK159167	CK159167 FGAS04056
41	289.5	18.5	1533	10 AG335255	AG335255 Mus muscu
42	289.5	18.5	1685	10 AG390959	AG390959 Mus muscu
43	289	18.5	1596	10 AG330370	AG330370 Mus muscu
44	288	18.4	787	3 BJ375707	BJ375707 BJ375707
45	287.5	18.4	681	3 BJ345323	BJ345323 BJ345323

ALIGNMENTS

RESULT 1
T09652 LOCUS 317 bp DNA linear GSS 05-DEC-2000
DEFINITION 0224m3 gmbpFHB3.1, G. Roman Reddy Plasmodium falciparum genomic
T09652
T09652.1 GI:319484
KEYWORDS GSS.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 317)
AUTHORS Reddy,G.R., Chakrabarti,D., Schuster,S.M., Ferl,R.J., Almira,E.C.
and Dame,J.B.
TITLE Gene sequence tags from Plasmodium falciparum genomic DNA fragments
JOURNAL prepared by the genebase activity of mung bean nuclease
PROC. Natl. Acad. Sci. U.S.A. 90, 9867-9871 (1993)
PUBMED 834327
COMMENT Other GSSes: 0224m7
Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.ufl.edu
Seq Primer: T3
Class: Shotgun.
Location/Qualifiers
1..317
/organism="Plasmodium falciparum"

FEATURES

source

/mol_type="genomic DNA"
/db_xref="taxon:5833"
/clone="0224m"
/lab_host="E. coli XL1-Blue"
/clone_1ib="gmbpFHB3.1, G. Roman Reddy"
/note="Vector: pBluescript SK(+); Genomic DNA, from
asynchronous blood stage parasites of the cloned Honduran
HB3 isolate cultured in vitro, was digested with mung bean
nuclease in the presence of 30% formamide at 500C
(Verneick, K.D., Imbercki, R.B., and McCutchan, T.F. 1988.
Nucleic Acids Research 16:6883-6896). The ends of the
fragments were ligated to EcoR V-cleaved and
dephosphorylated pBluescript SK(+). Recombinant plasmids
transformed E. coli XL1-Blue."

ORIGIN

Alignment Scores:
Pred. No.: 8,83e-31 Length: 317
Score: 396.00 Matches: 77
Percent Similarity: 96.25% Conservative: 0
Best Local Similarity: 96.25% Mismatches: 3
Query Match: 25.34% Indels: 0
DB: 10 Gaps: 0

US-09-837-344-31 (1-316) x T09652 (1-317)

QY 237 GIUULEPFOAIIIEGILEUPROSERGIUAENGUARGLYTYRTYRILEPROHISGIN 256
|||
DB 1 GAATTCGAAGCTTAGAACCTTCATCAGAAATGAACTGATATTATATACACATCA 60
QY 257 SerserleuProGluAspAsnArgIYanSerArgAspSerlysgluIeserlele 276
|||||
DB 61 TCTCTTACCTCAGACACACAGAGAAATAGTAGATTCCAGGAAATATCTATATA 120
QY 277 GIUYSThAsnArgGluSerIleThrThAsnValGluGluYArgAspIleHislys 296
|||||
DB 121 GAAAAAACAATAGAGAACTTATTCACAAATGTTGAAAGAGAAAGATATACATA 180
QY 297 GIUYSleuGluGluYAspArgIleSerIleYsPProGluGluYAspIlySer 316
|||||
DB 181 GGACATCTTGAAGAAAGAAAGATGTTCAATATAAACCAATACAAAGAAATTAATCT 240

RESULT 2

LOCUS BX200721 777 bp DNA linear GSS 13-MAR-2003
DEFINITION Danto rerio genomic clone DKEY-21101, genomic survey sequence.
ACCESSION BX200721
VERSION BX200721.1 GI:28032607
KEYWORDS GSS

SOURCE Danto rerio (zebrafish)
ORGANISM Danto rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danto.

REFERENCE 1 (bases 1 to 777)
AUTHORS Humphrey,S.J., Huckle,E. and Durham,J.L.
TITLE Direct Submision
JOURNAL Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphrey@sanger.ac.uk Unpublished
humphrey@sanger.ac.uk Unpublished

COMMENT This sequence was generated from the SP6 end of BAC 21101.21101 is
part of the Dantokey BAC library created by R. Plaetere and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/
Location/Qualifiers

FEATURES

source
1..777
/organism="Danto rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-21101"
/issue_type="Testis"
/note="Vector pindigobAC-536"

ORIGIN

Alignment Scores:
Pred. No.: 9,38e-30 Length: 777
Score: 390.50 Matches: 101
Percent Similarity: 66.41% Conservative: 69
Best Local Similarity: 39.45% Mismatches: 55
Query Match: 24.98% Indels: 31
DB: 10 Gaps: 12

US-09-837-344-31 (1-316) x BX200721 (1-777)

QY 3 LeuGluGluGluArgArgAlaYsGluYsleuGluGluGluGlnGlnSerAspleuGluGln 22
|||||
DB 764 CTTGAGAAAGAAAAAGAAAGAAAGAGAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAG 708
QY 23 AsparGleuAlaYsGluYsleuGluGlnGlnGlnSerAsp--LeuGluGlnGluArg 41
:::|
DB 707 GAGCCTCTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 648
QY 42 ArgAlaYsGluYsleuGluGlnGlnGlnSerAspleuGluGluGlnGluArgAlaYs 61
:::|
DB 647 CAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 594
QY 62 GIUYSleuGluGluGlnGlnSerAspleuGluGluGluGluArgAlaYsGluYsleu 81
:::|
DB 593 GAAAAA--AGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 537
QY 82 Gln-----GluGlnGlnSerAspleuGluGlnGlnSerAspleuAlaYsGluYs 97
:::|
DB 536 GAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 477
QY 98 LeuGln-----GluGlnGlnSerAspleuGluGluGluGluArgAlaYsGluYs 113
:::|
DB 476 CTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 417
QY 114 YsleuGln-----GluGlnGlnSerAspleuGluGluGluGluArgAlaYs 129
:::|
DB 416 CGTCTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 357
QY 130 GIUYSleuGluGluGlnGlnSerAsp-----LeuGluGlnGluArgAlaYs 146
:::|
DB 356 GAGCCTCTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 297
QY 147 GIUYSleuGln-----GluGlnGlnSerAspleuGluGluGluGluArgAla 162
:::|
DB 296 GAGCCCTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 237
QY 163 YsGluYsleu-----GluGluGlnGlnSerAspleuGluGluGluGluArgAlaYs 180
:::|
DB 236 AAGAGAAAGCTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 177
QY 181 GIUYSleuGln-----GluGlnGlnSerAspleuGluGluGluGluArgAlaYs 196
:::|
DB 176 GAGCCTCTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 117
QY 197 YsGluYsleuGln-----GluGlnGlnSerAspleuGluGluGluGluArgAlaYs 215
:::|
DB 116 AAGAGAAAGCTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 57
QY 216 YsAsnLeuGluArgAlaYsleuGluYsleuGluYsleuAlaYsleuAlaYs 231
:::|
DB 56 AAGCCTCTTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 9

RESULT 3

LOCUS CX031474 619 bp mRNA linear EST 03-JAN-2005
DEFINITION 1343594 NCCWA 10RT#3 Oncorhynchus mykiss cDNA 5', mRNA sequence.
ACCESSION CX031474
VERSION CX031474.1 GI:56975828
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss

QY 14 GINGINGInGInSerASePLeuGInGInaPArGLeuAlaLySGluLyLeuGInGIn 33
 Db 713 CAAGAAGAGCTATTGCAATTGAAGAGAAAGCTTCAAAAGAAATTGACAGAGAGAA 654
 QY 34 GInSeArSePLeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSeP 53
 Db 653 CGAAAAGCCCAAGAAAGAAAGATGAAGGTTAAAGAA-----GAGAGAGAGAGTTAGC 600
 QY 54 LeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSePLeuGIn 72
 Db 539 TTAGAGAGCGAAACAATCCAAAAGAAATTGAGAAAGAAAGAAAGCAAAAGAGAA 540
 QY 73 GInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSePLeuGInGInaPArG 92
 Db 539 GAGAAAGAAAGCAAAAGAGAAAGAAAGCAAAAGAGAAAGAGAAAGAGAAAGAA 483
 QY 93 LeuAlaLySGluLyLeuGInGInGInSeArSePLeuGInGInuArGArGAlaLyS 112
 Db 482 CAAGCAAAAGAGAGAA---GAGAAAGCAAGCAAAAGAGAAAGAAAGAAAGCAAGA 426
 QY 113 GInuLyLeuGInGInGInSeArSePLeuGInGInuArGArGAlaLySGluLyLeu 132
 Db 425 GAAAGAA---GCAGAAAGAAAGAGAGAAAGAGAAAGCAAAAGAAAGCAAAAGCTGAA 369
 QY 123 GInGInuGInGInSeArSePLeuGInGInuArGArGAlaLySGluLyLeuGInGIn 152
 Db 368 GAAAGAGAGAAATTAAATTGAAGAGAAAGAAAGCAAAAGAGAAAGAAAGAAAGAA 309
 QY 153 GInSeArSePLeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSeP 172
 Db 308 GCAAAAGAACTTGAAGAGAAAGAAAGCAAAAGAAAGCAAAAGAGAGAAAGAAATTAA 249
 QY 173 LeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSePLeuGIn 192
 Db 248 TTAGAGAAAGAAAGCTTGAAGAAAGAAAGCAAAAGAAAGAAAGAAAGAAAGAAAGAA 189
 QY 193 GInuArGArGAlaLySGluLyLeuGInGInGInuArGArGAlaLySGluLyLeuGInGIn 212
 Db 188 GAAAGAGAGTTAAATTGAAGAGAGAGAAAGCTTCAAAAGAACTTGAAGAGAGAGAAAGAA 129
 QY 213 AArPThLyLeuASePLeuGInuArGArGAlaLySGluLyLeuGInGInSeArSePLeu 232
 Db 128 GAAAGAAAGAAAGAGAGCTTGAATATGATTTATGATGATGATGATGATGATGATGATGAT 69
 QY 233 TArGArGArGInGInGInSeArSePLeuGInGInGInSeArSePLeuGInGInGIn 252
 Db 68 AATGGGCTAGAAAGAAATTGAAGGCTAGAGAAAGAGCTCAGTGAAGAAAGAGAT-----TAT 15
 QY 253 TLePAr 254
 Db 14 CTTCA 9
 RESULT 5
 C2247380 743 bp DNA linear GSS 11-FEB-2005
 LOCUS A1AA-aef35a12.b1 Ancylosetoma caninum whole genome shotgun library
 DEFINITION (A1AAGS 001) Ancylosetoma caninum genomic, genomic survey sequence.
 ACCESSION C2247380
 VERSION C2247380.1 GI:59632821
 KEYWORDS GSS.
 SOURCE Ancylosetoma caninum (dog hookworm)
 ORGANISM Ancylosetoma caninum
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongyloida;
 Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
 1 (bases 1 to 743)
 REFERENCE Miteva, M., McCarter, J.P., Pape, D., Ritter, E., Tsagaris, V.I., R.,
 AUTHORS Ronko, I., Martin, J., Wylie, T., Dante, M., Meyer, R., Messina, D.,
 Watson, R.H., Clifton, S.W. and Wilson, R.,
 TITLE Genome Survey sequences from the parasitic nematode Ancylostoma
 caninum
 JOURNAL Unpublished (2004)
 COMMENT Contact: Miteva M

Washington University in St. Louis
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: nematode@wustl.edu
 Genomic DNA provided by John Hawdon (mcmjmh@wumc.edu) DNA
 sequenced by Washington University Genome Sequencing Center
 Class: shotgun.

FEATURES
 source
 Location/Qualifiers
 1..743
 /organism="Ancylosetoma caninum"
 /mol_type="genomic DNA"
 /strain="Baltimore"
 /db_xref="taxon:29170"
 /dev_stage="Adult"
 /lab_host="GSI0"
 /clone_lib="Ancylostoma caninum whole genome shotgun
 library (A1AAGS 001)"
 /note="Vector: pOTW13; Site 1: BstXI; Site 2: BstXI;
 Ancylostoma caninum genomic DNA was randomly sheared,
 end-repaired and size fractionated to enrich for 2-4 kb
 fragments. Genomic DNA was provided by John Hawdon
 (mcmjmh@wumc.edu) at George Washington University.
 Sequencing by Washington University Genome Sequencing
 Center, St. Louis, MO."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,926-23 Length: 743
 Score: 328.50 Matches: 59
 Percent Similarity: 65.32% Conservative: 86
 Best Local Similarity: 26.58% Mismatches: 74
 Query Match: 21.02% Indels: 3
 DB: 10 Gaps: 2

US-09-837-344-31 (1-316) x C2247380 (1-743)
 QY 4 GINGINGInGInSerASePLeuGInGInaPArGLeuAlaLySGluLyLeuGInGIn 23
 Db 17 CAAGAAGAGCTATTGCAATTGAAGAGAAAGCTTCAAAAGAAATTGACAGAGAGAA 76
 QY 24 GInSeArSePLeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSeP 41
 Db 77 CGAAAAGCCCAAGAAAGAAAGATGAAGGTTAAAGAA-----GAGAGAGAGAGTTAGC 136
 QY 42 LeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSePLeuGIn 61
 Db 137 CGAAAAGCCCAAGAAAGAAAGATGAAGGTTAAAGAA-----GAGAGAGAGAGTTAGC 196
 QY 62 LeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSePLeuGIn 81
 Db 197 CGAAAAGCCCAAGAAAGAAAGATGAAGGTTAAAGAA-----GAGAGAGAGAGTTAGC 256
 QY 82 GInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSePLeuGInGInaPArG 101
 Db 257 GAGAAAGAAAGCAAAAGAGAAAGAAAGCAAAAGAGAAAGAGAAAGAGAAAGAGAAAGAA 316
 QY 102 GInSeArSePLeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSeP 121
 Db 317 CGAAAAGCCCAAGAAAGAAAGATGAAGGTTAAAGAA-----GAGAGAGAGAGTTAGC 376
 QY 122 LeuGInGInGInuArGArGAlaLySGluLyLeuGInGInGInSeArSePLeuGIn 141
 Db 377 CGAAAAGCCCAAGAAAGAAAGATGAAGGTTAAAGAA-----GAGAGAGAGAGTTAGC 436
 QY 142 GInuArGArGAlaLySGluLyLeuGInGInGInSeArSePLeuGInGInuArGArG 161
 Db 437 CAAGCAAAAGAGAGAA---GAGAAAGCAAGCAAAAGAGAAAGAGAAAGAGAAAGAGAAAGAA 496
 QY 162 GInuLyLeuGInGInGInSeArSePLeuGInGInuArGArGAlaLySGluLyLeu 181
 Db 497 CAAGCAAAAGAGAGAGAA---GAGAAAGCAAGCAAAAGAGAGAAAGAGAAAGAGAAAGAGAA 556


```
Db      15 AGAGAGCGCTGAAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGCT 74
Qy      27 LysGluLysLeu---GlnGluGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGln 45
Db      75 AAAAGAAAAATTGAAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 134
Qy      46 LysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 65
Db      135 -----GAAAGAAACAAAGAAAGCGCTTGAAGAAAG-----AAAAATTAATTGAA 179
Qy      66 GlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 85
Db      180 GAAAGAAACAAAGAAAGATTAGAGAAAGAAAGAAAG-----AAGCGACAAAGAAAGAA 230
Qy      86 SerAspLeuGlnGlnSerAspLeuAlaLysGluLysLeuGlnGlnGlnGlnGln 105
Db      231 AAACGCTTGAAGCGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 287
Qy      106 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
Db      288 GAGGAAAGAAAGAAAGAAAGAAAGAAAGCGCTAGAGAAAGAAAGAAAGAAAGCGTAA 347
Qy      121 AspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
Db      348 AAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 407
Qy      141 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 155
Db      408 GAAAGCGCTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCGTAA 467
Qy      156 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 175
Db      468 GCTGAGCGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 524
Qy      176 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 191
Db      525 GAAAGCTGAAAGAAAGAAAGAAAGAAAGCGTAAAGCGGAAAGAAAGAAAGAAAGAA 584
Qy      192 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 211
Db      585 GAGGAGCTTAAAGAAAGAAAGAAAGCGGCAAGAAAGAAAGAAAGAAAGCTTAAAGAAAGAAAC 644
Qy      212 AlaAspThrLysLeuAsnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 231
Db      645 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 704
Qy      232 LeuTyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 251
Db      705 GAAAAAAACGT-----AAAGCCGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 740
Qy      252 TyrTleProHisGlnSerSerLeuProGlnAspAsnArgGlnAsnSerArgAspSerLys 271
Db      741 AAAGAAAGAAAGAAAGAAAGCGGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 800
Qy      272 GlnLysSerLysLysLeu-----LysThrAsnArgLysSerLysLysLysLys 286
Db      801 GAAAGCTAAAGTACAGATGAAGAAAGAAAGCTGAAAGAAAGAAAGAAAGAAAGCTTTAGAGTT 860
Qy      287 AsnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 303
Db      861 TAAACAGAA-----GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 881

RESULT 10
A2536100      868 bp      DNA      linear      GSS 03-NOV-2000
LOCUS
DEFINITION
ENTCO15TF Entamoeba histolytica Shared DNA Entamoeba histolytica
genomic, genomic survey sequence.
ACCESSION
A2536100
VERSION
A2536100.1 GI:11093047
KEYWORDS
GSS.
SOURCE
Entamoeba histolytica
Entamoeba histolytica
ORGANISM
Eukaryota; Entamoebidae; Entamoeba.
```

```
REFERENCE
1 (bases 1 to 868)
AUTHORS
Loftus,B., Van Aken,S. and Fraser,C.
TITLE
Determination of clone end sequences from Entamoeba histolytica
JOURNAL
HM1:IMSS sheared DNA library
COMMENT
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 74
High quality sequence stop: 837.
FEATURES
source
1..868
location/Qualifiers
1..868
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica sheared DNA"
/notes="Vector: pHOSt1; Site 1: Bat I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaubin and B.
Bartell, Oxford University Press, 1999)."
```

ORIGIN

Alignment Scores:

Pred. No.:	1..926-22	Length:	868
Score:	319.50	Matches:	93
Percent Similarity:	52.81%	Conservative:	76
Best Local Similarity:	29.06%	Mismatches:	88
Query Match:	20.44%	Indels:	63
DB:	9	Gaps:	13

US-09-837-344-31 (1-316) x A2536100 (1-868)

```
Qy      4 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 23
Db      29 GAGGCTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 82
Qy      24 ArgLeuAlaLysGluLysLeu---GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 42
Db      83 -----GCTAAAGAAAGAAATTTGACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 136
Qy      43 AlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 62
Db      137 AAAAAAGAA-----GAAAGAAAGAAAGAAAGCGCTTGAAGAAAG-----AAAAAA 181
Qy      63 LysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 82
Db      182 TTAATTGGAAGAAAGAAAGAAAGATTAGAGAAAGAAAGAAAG-----AAAGCAGAA 232
Qy      83 GlnGlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 102
Db      233 GAAAGAAAGAAAGCGCTTGAAGCTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 289
Qy      103 SerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
Db      290 AAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCGCTGAGAAAGAAAGAAAGAA 349
```

Qy 118 GlnGlnSerAsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnSer 137
Db 350 GAACGTAAAGAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 409
Qy 138 AsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 152
Db 410 TTTAAAGAAAGAAAGCGCTAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 469
Qy 153 GlnSerAsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 172
Db 470 AAACGTAAAGCTGAGCGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 526
Qy 173 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 188
Db 527 GAAGAAAGAAAGAAAGCGCTAGAAAGAAAGAAAGAAAGAAAGAAAG 586
Qy 189 AsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 208
Db 587 AGAAAGAGAGAGGAGTTTAAAGAAAGAAAGAGAGAGAGAGAGAGAGAG 646
Qy 209 GlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 228
Db 647 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 694
Qy 229 AlaGlnAsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 248
Db 695 -----GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 706
Qy 249 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 268
Db 707 CGA-----CCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 736
Qy 269 AsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 288
Db 737 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 796
Qy 289 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 304
Db 797 ACCGCGAGAAAGAAAGCGTTGAGAAAGAAAGAAAGAAAGAAAGAAAG 856

RESULT 11
CN036384 771 bp mRNA linear EST 29-MAR-2004
LOCUS CN036384
DEFINITION nm.15.c6.c3_Mach Ambystoma mexicanum cDNA, mRNA sequence.
ACCESSION CN036384
VERSION CN036384.1 GI:45806755
KEYWORDS EST.
SOURCE Ambystoma mexicanum (axolotl)
ORGANISM Ambystoma mexicanum
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
Ambystoma.
1 (bases 1 to 771)
Putte, S., Smith, J.J., Walker, J.A., Rondelet, M., Westrock, D.,
Morgan, J., Samuels, A.K., Kump, K., King, D.C., Maness, N.J.,
Haberman, B., Tanaka, E., Bryant, S.V., Gardiner, D.M., Parichy, D.M.
and Voss, S.R.
From biomedicine to natural history research: EST resources for
Ambystoma mexicanum
BMC Genomics 5 (1), 54 (2004)
15310388
Contact: SR Voss
Department of Biology
University of Kentucky
TH Morgan Building, Lexington, KY 40506, USA
Tel: 859 257 9888
Fax: 859 257 1717
Email: svoss@uky.edu
The EST is quality trimmed at the ends with a 20 base window and
quality threshold of 15 (phred quality score). Please visit
http://balamander.uky.edu for any information (trace, quality files
etc) regarding this EST.

FEATURES
source Location/Qualifiers
1..771
/organism="Ambystoma mexicanum"
/mol_type="mRNA"
/db_xref="taxon:8296"
/issue_type="Limb Blastema and Proximal Limb Tissue
collected from larvae on days 1-6 of regeneration"
/clone_lib="Mach"
ORIGIN
Alignment Scores:
Pred. No.: 1.87e-22 Length: 771
Score: 319.00 Matches: 74
Percent Similarity: 66.09% Conservative: 78
Best Local Similarity: 32.17% Mismatches: 66
Query Match: 20.41% Indels: 12
DB: 7 Gaps: 8
US-09-837-344-31 (1-316) x CN036384 (1-771)
Qy 2 AsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 21
Db 89 GATTGTGAGCGTGAAGATAGACGGAAGAAAGAAAGAAAGAAAGAAAGAAAG 148
Qy 22 GlnAsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
Db 149 GATGAGAAAGAAAGAAAGCGGAGCGGAAAGGAAAGAAAGAAAGAAAGAAAG 208
Qy 41 ArgArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
Db 209 CGT-----GACGCGGTCCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 256
Qy 61 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
Db 257 AGAGAGCGGAGAGAGCGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
Qy 81 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
Db 317 GAGCGTGAAGAGAGCGCTGAGCGTGAAGAGAGAGAGAGAGAGAGAGAG 373
Qy 101 GlnGlnSerAsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
Db 374 AAGAGCGGAGAGCGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
Qy 118 GlnGlnSerAsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 137
Db 434 AGCGAGAGAGAGCGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 493
Qy 138 AsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 157
Db 494 GAGCGTGAAGAGAGAGCGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 553
Qy 158 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 177
Db 554 GAAGAAAGAAAGAGCGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 613
Qy 178 ArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
Db 614 GAA-----AGAGAAAGAGAGCGCTGAAGAGAGAGAGAGAGAGAGAGAG 670
Qy 197 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 215
Db 671 AAGGAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730
Qy 216 GlnAsepleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 225
Db 731 AAGGAGCGAG 760

RESULT 12
A2337339/c 650 bp DNA linear GSS 29-SEP-2000
LOCUS A2337339
DEFINITION IM0068B03F Mouse 10kb plasmid UUCGIM library Mus musculus genomic
clone UUCGIM0068B03 F, genomic survey sequence.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:28:45 ; Search time 7029.36 Seconds
(without alignments)
2555.359 Million cell updates/sec

Title: US-09-837-344-31
Perfect score: 1563
Sequence: 1 SDBEQERRAKEKIQEQOSD.....GHLEKKDGSIKPEQKEDKS 316

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Rgapop 6.0 , Rgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O/cgcn2.1/USPTO.spool.h/US09837344/runat.05122005.133532.734/app.query.faeta_1.718
-DB=GenEmbl -QMT=faetap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344 @CGN 1.1 5142 @runat.05122005.133532.734 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MANY TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -RGAPOP=6
-RGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_pro:*
10: gb_rcv:*
11: gb_by:*
12: gb_un:*
13: gb_vl:*
14: gb_hlg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	99.1	950	6	A28740
2	1549	99.1	950	6	ARI62961
3	1549	99.1	950	6	ARI78250

4	1549	99.1	1482	6	ARI62970	ARI62970 Sequence
5	1549	99.1	1482	6	ARI78259	ARI78259 Sequence
6	1549	99.1	1482	6	PELSA13	230320 P.falciparu
7	1546	98.9	1482	6	ARI62974	ARI62974 Sequence
8	1546	98.9	1482	6	ARI78263	ARI78263 Sequence
9	1494	95.6	5970	2	PELSA1G	X56203 P.falciparu
10	1465	93.7	1493	6	A28743	A28743 cDNA for IS
11	1465	93.7	1493	6	ARI62967	ARI62967 Sequence
12	1465	93.7	1493	6	ARI78256	ARI78256 Sequence
13	1453.5	93.0	253001	2	AE014834	AE014834 Plasmodiu
14	771.5	49.4	954	6	ARI62966	ARI62966 Sequence
15	771.5	49.4	954	6	ARI78255	ARI78255 Sequence
16	771.5	49.4	988	6	PELSA15	230319 P.falciparu
17	771.5	49.4	988	6	A28742	A28742 cDNA for IS
18	771.5	49.4	988	6	ARI62963	ARI62963 Sequence
19	771.5	49.4	988	6	ARI78252	ARI78252 Sequence
20	740.5	47.4	1368	11	AY751501	AY751501 Synthetic
21	554	35.4	843	2	PELSA1C	L40885 Plasmodium
22	554	35.4	843	2	PELSA1H	L40890 Plasmodium
23	554	35.4	843	2	PELSA1I	L40891 Plasmodium
24	553	35.4	843	2	PELSA1B	L40887 Plasmodium
25	553	35.4	843	2	PELSA1E	L40887 Plasmodium
26	553	35.4	843	2	PELSA1F	L40888 Plasmodium
27	553	35.4	843	2	PELSA1K	L40893 Plasmodium
28	550	35.2	843	2	PELSA1G	L40889 Plasmodium
29	550	35.2	843	2	PELSA1A	L40834 Plasmodium
30	549	35.1	843	2	PELSA1B	L40909 Plasmodium
31	549	35.1	843	2	PELSA1C	L40910 Plasmodium
32	549	35.1	843	2	PELSA1D	L40947 Plasmodium
33	549	35.1	843	2	PELSA1N	L40886 Plasmodium
34	549	35.1	843	2	PELSA1D	L40892 Plasmodium
35	549	35.1	843	2	PELSA1J	L40835 Plasmodium
36	549	35.1	843	2	PELSA1M	L40837 Plasmodium
37	549	35.1	843	2	PELSA1O	L40837 Plasmodium
38	549	35.1	843	2	PELSA1O	L40837 Plasmodium
39	469	30.0	795	2	AF246996	AF246996 Plasmodium
40	423	27.1	170771	5	BX465202	BX465202 Zebrafish
41	420.5	26.9	194351	5	AL772295	AL772295 Zebrafish
42	420.5	26.9	205118	5	AL928674	AL928674 Zebrafish
43	408	26.1	214506	14	AC159753	AC159753 Bos tauru
44	407.5	26.1	220046	9	AC124587	AC124587 Mus muscu
45	397	25.4	82400	9	AC090495	AC090495 Genomic s

ALIGNMENTS

RESULT 1					
A28740	A28740	950 bp	DNA	linear	PAT 04-JUN-1995
LOCUS					
DEFINITION	CDNA for LSA-R-NR protein.				
ACCESSION	A28740				
VERSION	A28740.1	GI:1247512			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
	Hominidae; Homo.				
REFERENCE	1 (bases 1 to 950)				
AUTHORS					
TITLE	LIVER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING				
JOURNAL	EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES				
FEATURES	Patent: WO 9213884-A 29 20-AUG-1992;				
source	Location/Qualifiers				
	1..950				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN
Alignment Scores: 1.01e-85 Length: 950
Pred. No.: 1549.00 Matches: 314
Score:

Percent Similarity: 99.37%
Best Local Similarity: 99.37%
Query Match: 99.10%
DB: 6
Gaps: 0

US-09-837-344-31 (1-316) x A28740 (1-950)

```
QY 1 SerAapLeuGluGlnGluArgAlaLysGluLysLeuGlnGlnSerAapLeu 20
DB 3 AGCGATCTTAGAACAAAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 62
QY 21 GlnGlnAapArgLeuAlaLysGluLysLeuGlnGlnGlnSerAapLeuGlnGln 40
DB 63 GAACAAGATAGACTTCTTAAGAAAGTTGCAAGAGCAAGCAAGGATTTAACAAGAG 122
QY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnSerAapLeuGlnGlnArgAla 60
DB 123 AGACTTGCTTAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGACGCTGCT 182
QY 61 LysGluLysLeuGlnGlnGlnSerAapLeuGlnGlnArgAlaLysGluLys 80
DB 183 AAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGACGCTTAAGAAAG 242
QY 81 LeuGlnGlnGlnGlnSerAapLeuGlnGlnAapArgLeuAlaLysGluLysLeuGlnGln 100
DB 243 TTGCAAGAACAAAGCGATTTAAGAACAAAGCGATTTAAGAAAGTTTAACAAGAG 302
QY 101 GlnGlnSerAapLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnSer 120
DB 303 CAGCAAAAGCGATTTAAGAACAAAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGC 362
QY 121 AapLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnSerAapLeuGln 140
DB 363 GATTTAAGAACAAAGAGACGCTGCTTAAGAAAGTTGCAAGAACAAAGCGATTTAAGAA 422
QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnGln 160
DB 423 CAAAGAGACGCTTAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGAGAG 482
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnArgAlaLys 180
DB 483 CCGTCTTAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGACGCTGCTAA 542
QY 181 GlnLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnGlnArgAlaLysGluLysLeu 200
DB 543 GAAGAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGACGCTTAAGAAAGGTTG 602
QY 201 GlnGlnGlnGlnArgAapLeuGlnGlnArgAlaLysAlaAapThrLysLysAapLeuGln 220
DB 603 CAAAGAGCAAGAAAGATTTAAGAACAAAGAGCGCTGATACGAAAGAAATTTAGAAAG 662
QY 221 LysLysGluLysGluLysAlaLysAlaGluAapLeuThrLysArgLeuGlnLysProAla 240
DB 663 AAAAGAGCAATGCAATATATATATGCAAGAGATTATATGCTTAAAGAAATCAACGCT 722
QY 241 LysGluLysProSerGluLysGlnGlnArgAlaLysAlaLysAlaLysAlaLysAlaLys 260
DB 723 ATAGAACTTCATCAAGAAATGAAAGCGATTAATATACACATCAATCTTCTTTAAGCT 782
QY 261 GlnAapAapArgGluLysAapSerArgAapSerLysGluLysLysLysLysLysLys 280
DB 783 CAGAGCAACAGAGGAAATGAGATTCACAAAGAAATATCTATATAGAAAGAAACAAAT 842
QY 281 ArgGluSerLysThrAapAapAlaGluLysArgAapLysLysLysGluLysLysGlu 300
DB 843 AGAGATCTATTAACAATATGTAAGAGCGAAGGAGATATACATTAAGAGCACTTTGAA 902
QY 301 GlnLysLysAapGluLysLysLysProGlnGlnLysGlnAapLysSer 316
DB 903 GAAGAAAGAGATGCTCAATTAAGAACAAAGCAAGAAAGAGATTAATCT 950
```

RESULT 2
AR162961

```
LOCUS AR162961 950 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 32 from patent US 6270771.
ACCESSION AR162961
VERSION AR162961.1 GI:16233425
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 950)
AUTHORS Guerin-Marchand,C. and Druilhé,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
JOURNAL bearing epitopes capable of stimulating the T lymphocytes
FEATURES
source location/Qualifiers
1..950
/organism="Unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1,01e-85 Length: 950
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 6 Gaps: 0
US-09-837-344-31 (1-316) x AR162961 (1-950)
QY 1 SerAapLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnSerAapLeu 20
DB 3 AGCGATCTTAGAACAAAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 62
QY 21 GlnGlnAapArgLeuAlaLysGluLysLeuGlnGlnGlnSerAapLeuGlnGln 40
DB 63 GAACAAGATAGACTTCTTAAGAAAGTTGCAAGAGCAAGCAAGGATTTAACAAGAG 122
QY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnArgAla 60
DB 123 AGACTTGCTTAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAAAGAGACGCTGCT 182
QY 61 LysGluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnGlnArgAlaLysGluLys 80
DB 183 AAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGACGCTTAAGAAAG 242
QY 81 LeuGlnGlnGlnGlnSerAapLeuGlnGlnAapArgLeuAlaLysGluLysLeuGlnGln 100
DB 243 TTGCAAGAACAAAGCGATTTAAGAACAAAGCGATTTAAGAAAGTTTAACAAGAG 302
QY 101 GlnGlnSerAapLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnSer 120
DB 303 CAGCAAAAGCGATTTAAGAACAAAGAGACGCTGATTAAGAAAGTTGCAAGAACAAAGC 362
QY 121 AapLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnSerAapLeuGln 140
DB 363 GATTTAAGAACAAAGAGACGCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTAAGAA 422
QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnGln 160
DB 423 CAAAGAGACGCTTAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGAGAG 482
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnArgAlaLys 180
DB 483 CCGTCTTAAGAAAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGACGCTGCTAA 542
QY 181 GlnLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnArgAlaLysAlaAapThrLys 200
DB 543 GAAGAGTTGCAAGAACAAAGCGATTTAAGAACAAAGAGAGACGCTTAAGAAAGGTTG 602
QY 201 GlnGlnGlnGlnArgAapLeuGlnGlnArgAlaLysAlaAapThrLysLysAapLeuGln 220
DB 603 CAAAGAGCAAGAAAGATTTAAGAACAAAGAGCGATTAAGAAAGAAATTTAGAAAG 662
```

Qy 221 LysLysGluuHISGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIleProAla 240
Db 663 AAAAAGGAACATGAGATATATGACAGAGATTTATATGATCGTTTAAATTCAGACT 722
Qy 241 ILeGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHISGlnSerSerLeuPro 260
Db 723 ATAAACTTCATCAGAAAATGAAACGTGATATATATACCAATCATCTCTTTACCT 782
Qy 261 GlnAspAsnArgGlyAsnSerArgAspSerGlyLeuIleSerIleIleGluLeuThrAsn 280
Db 783 CAGACAACAAGAGGAAATAGTAGAGATTCCAAAGAAATATCTATATATGAAAAACAAAT 842
Qy 281 ArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHISLysGlyHISLeuGlu 300
Db 843 AGAAGATCTATTACAAATGTTGAGACGAAGGATATACATTAAGACATCTTGAA 902
Qy 301 GlnLysLysAspGlySerIleLysProGluLysGluAspLysSer 316
Db 903 GAAAAGAAAGATGTTCAATTAACCAAGACAAAGAAAGATTAATCT 950

RESULT 3
ARI78250 950 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 32 from patent US 6319502.
DEFINITION ARI78250
ACCESSION ARI78250 GI:20219388
VERSION ARI78250.1 GI:20219388
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
AUTHORS 1 (bases 1 to 950)
TITLE Guerlin-Marchand,C. and Druille,P.
JOURNAL Peptide sequences specific for the hepatic stages of P. falciparum
FEATURES bearing epitopes capable of stimulating the T lymphocytes
SOURCE Patent: US 6319502-A 32-20-NOV-2001;
1. 950 Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.01e-85 Length: 950
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 6 Gaps: 0

US-09-837-344-31 (1-316) x ARI78250 (1-950)

Qy 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeu 20
Db 3 AGCGATCTTAGAACAAAGAGACGCTCTTAAAGAAAGTTGCAAGAACAAAGCATTTA 62
Qy 21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlu 40
Db 63 GAACAAATAGACTGCTTAAAGAAAGTTTACAAAGACGCAAAAGCATTTTGAACAAGAG 122
Qy 41 ArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAla 60
Db 123 AGACTTGTAAAGAAAGGTTGTCMAAGAACAAAGCATCTTAGAACAAAGAGACGCTCT 182
Qy 61 LysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAlaLysGluLys 80
Db 183 AAAAGAAAAGTTGCAAGAACAAAGCATTTTGAACAAGAGACGCTTAAAGAAAAG 242
Qy 81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGlu 100
Db 243 TTGCAAGAACAAAGCATTTTGAACAAGATAGACTTGTAAAGAAAAGTTTACAAAGAG 302
Qy 101 GlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 303 CAGCAAGCATTTTAGAACAAAGAGACGCTCTTAAAGAAAAGTTGCAAGAACAAAGC 362
Qy 121 AspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
Db 363 GATTTAGAACAAAGAGACGCTCTTAAAGAAAAGTTGCAAGAACAAAGCATTTTGA 422
Qy 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlu 160
Db 423 CAAAGAGACTTCTTAAAGAAAAGTTGCAAGAACAAAGCATTTTGAACAAGAGAG 482
Qy 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnSerAspLeuGlnGlnGluArgAlaLys 180
Db 483 CGTCTTAAAGAAAAGTTGCAAGAACAAAGCATTTTGAACAAGAGACGCTTAA 542
Qy 181 GlnLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 543 GAAAAGTTGCAAGAACAAAGCATTTTGAACAAGAGACGCTTAAAGAAAAGCTTG 602
Qy 201 GlnGlnGlnGlnGlnArgAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
Db 603 CAAAGACGCAAGAGATTTTGAACAAGAGACGCTCTTGAAGAAAAGTTTGAAGA 662
Qy 221 LysLysGluuHISGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIleProAla 240
Db 663 AAAAAGGAACATGAGATATATAGCAGAGATTTATATGATCGTTTAAATTCAGACT 722
Qy 241 ILeGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHISGlnSerSerLeuPro 260
Db 723 ATAGAACCTTCATCAGAAAATGAAACGTGATATATATACACATCTCTTTACCT 782
Qy 261 GlnAspAsnArgGlyAsnSerArgAspSerGlyLeuIleSerIleIleGluLeuThrAsn 280
Db 783 CAGACAACAAGAGGAAATAGTAGAGATTTCCAAAGAAATATCTATATAGAAAAACAAAT 842
Qy 281 ArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHISLysGlyHISLeuGlu 300
Db 843 AGAAGATCTATTACAAATGTTGAGAGACGAAGGATATACATTAAGACATCTTGAA 902
Qy 301 GlnLysLysAspGlySerIleLysProGluLysGluAspLysSer 316
Db 903 GAAAAGAAAGATGTTCAATTAACCAAGACAAAGAAAGATTAATCT 950

RESULT 4
ARI62970 1482 bp DNA linear PAT 17-OCT-2001
LOCUS Sequence 41 from patent US 6270771.
DEFINITION ARI62970
ACCESSION ARI62970
VERSION ARI62970.1 GI:16233435
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
AUTHORS 1 (bases 1 to 1482)
TITLE Guerlin-Marchand,C. and Druille,P.
JOURNAL Peptide sequences specific for the hepatic stages of P. falciparum
FEATURES bearing epitopes capable of stimulating the T lymphocytes
SOURCE Patent: US 6270771-A 41-07-AUG-2001;
1. 1482 Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.52e-85 Length: 1482
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 6 Gaps: 0

US-09-837-344-31 (1-316) x ARI62970 (1-1482)

QY 1 SerApeleuglunGluInaArgAAlaYseGluYleuGlnGlnInSerApeleu 20
Db 13 AGCGATCTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAGAACACAAAGCATTTA 72
QY 21 GluInApeAArgLeuAlaYseGluYleuGlnGlnInSerApeleuGlnGlu 40
Db 73 GAAACAAGATAGACTTGCTTAAGAAAAGTTACAAAGAGCAGAAACGATTTAGAACAAAG 132
QY 41 ArgAgaAlaYseGluYleuGlnGlnInSerApeleuGlnGlnInaArgAAla 60
Db 133 AGACTTGCTTAAGAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCT 192
QY 61 LysGluYleuGlnGlnInSerApeleuGlnGlnInaArgAAlaYseGluYs 80
Db 193 AAAAGAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCTTAAGAAAAG 252
QY 81 LeuGlnGlnGlnInSerApeleuGlnInApeAArgLeuAlaYseGluYleuGlnGlu 100
Db 253 TTGCAGAACACAAAGCATTTAGAACAGATGACTTGCTTAAGAAAAGTTACAAAG 312
QY 101 GlnGlnInSerApeleuGlnGlnInaArgAAlaYseGluYleuGlnGlnInSer 120
Db 313 CAGCAAGCGATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAGAACACAAAGC 372
QY 121 ApeleuGlnGlnInaArgAAlaYseGluYleuGlnGlnInSerApeleuGlu 140
Db 373 GATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAGAACACAAAGCATTTAGAA 432
QY 141 GlnGlnInaArgAAlaYseGluYleuGlnGlnInSerApeleuGlnGlnInaArg 160
Db 433 CAAGAGAGACTTGCTTAAGAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAG 492
QY 161 ArgAAlaYseGluYleuGlnGlnInSerApeleuGlnGlnInaArgAAlaYs 180
Db 493 CGTGCTTAAGAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCTTAAG 552
QY 181 GluYleuGlnGlnInSerApeleuGlnGlnInaArgAAlaYseGluYleu 200
Db 553 GAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCTTAAGAAAAGTTG 612
QY 201 GlnGlnInaArgAAlaYseGlnGlnInaArgAAlaYseGlnGlnInaArg 220
Db 613 CAAGAGAGCATTTAGAACAGAGAGCGTCTTAAGAAAAGTTAGAAAAGAG 672
QY 221 LysYleuGlnGlnInaArgAAlaYseGlnGlnInaArgAAlaYseGlnGlnInaArg 240
Db 673 AAAAGGAAAGATGAGATATTATAGAGAGATTTATATGCTTTAGAAAATACCACT 732
QY 241 IleGlnLeuProSerGlnAseGlnGlnInaArgAAlaYseGlnGlnInaArg 260
Db 733 ATAGAACTTCATCAGAAAATGAGCGTATATTATATCAGATCATCTTTTACCT 792
QY 261 GlnApeAArgGlnYseGlnGlnInaArgAAlaYseGlnGlnInaArg 280
Db 793 CAGGAGACAGAGGAGATAGATTCAGAGAAATATTATATATGAAAACAAAT 852
QY 281 ArgGlnSerIleThrThrAsnValGlnGlnInaArgAAlaYseGlnGlnInaArg 300
Db 853 AAGAGATCTTATTAACAACAATGTTGAGAGAGAGGATATATTAAGAGACATCTTGAA 912
QY 301 GluYleuYseGlnGlnInaArgAAlaYseGlnGlnInaArgAAlaYseGlnGln 316
Db 913 GAAAGGAAAGATGCTTCAATTAACAGAACAAAGGAAAGATTAATCT 960

RESULT 5
ARI78259 1482 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 41 from patent US 6319502.
DEFINITION
ACCESSION ARI78259
VERSION ARI78259.1 GI:20219397
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
REFERENCE 1 (bases 1 to 1482)
AUTHORS Guerin-Merchand,C. and Druille,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
JOURNAL bearing epitopes capable of stimulating the T lymphocytes
Patent: US 6319502-A 41 20-NOV-2001;
FEATURES
source
1..1482
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.52e-85 Length: 1482
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
Gaps: 0
DB: 6
US-09-837-344-31 (1-316) x ARI78259 (1-1482)
QY 1 SerApeleuglunGluInaArgAAlaYseGluYleuGlnGlnInSerApeleu 20
Db 13 AGCGATCTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAGAACACAAAGCATTTA 72
QY 21 GluInApeAArgLeuAlaYseGluYleuGlnGlnInSerApeleuGlnGlu 40
Db 73 GAAACAAGATAGACTTGCTTAAGAAAAGTTACAAAGAGCAGAAACGATTTAGAACAAAG 132
QY 41 ArgAgaAlaYseGluYleuGlnGlnInSerApeleuGlnGlnInaArgAAla 60
Db 133 AGACTTGCTTAAGAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCT 192
QY 61 LysGluYleuGlnGlnInSerApeleuGlnGlnInaArgAAlaYseGluYs 80
Db 193 AAAAGAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCTTAAGAAAAG 252
QY 81 LeuGlnGlnGlnInSerApeleuGlnInApeAArgLeuAlaYseGluYleuGlnGlu 100
Db 193 GAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCTTAAGAAAAGTTG 252
QY 101 GlnGlnInSerApeleuGlnGlnInaArgAAlaYseGluYleuGlnGlnInSer 120
Db 253 TTGCAGAACACAAAGCATTTAGAACAGATGACTTGCTTAAGAAAAGTTACAAAG 312
QY 121 ApeleuGlnGlnInaArgAAlaYseGluYleuGlnGlnInSerApeleuGlu 140
Db 373 GATTTAGAACAGAGAGCGTCTTAAGAAAAGTTGCAGAACACAAAGCATTTAGAA 432
QY 141 GlnGlnInaArgAAlaYseGluYleuGlnGlnInSerApeleuGlnGlnInaArg 160
Db 433 CAAGGAGACTTGCTTAAGAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAG 492
QY 161 ArgAAlaYseGluYleuGlnGlnInSerApeleuGlnGlnInaArgAAlaYs 180
Db 493 CGTGCTTAAGAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCTTAAG 552
QY 181 GluYleuGlnGlnInGlnInSerApeleuGlnGlnInaArgAAlaYseGluYleu 200
Db 553 GAAAAGTTGCAGAACACAAAGCATTTAGAACAGAGAGCGTCTTAAGAAAAGTTG 612
QY 201 GlnGlnGlnInaArgAAlaYseGlnGlnInaArgAAlaYseGlnGlnInaArg 220
Db 613 CAAGAGAGCATTTAGAACAGAGAGCGTCTTAAGAAAAGTTAGAAAAGAG 672
QY 221 LysYleuGlnGlnInaArgAAlaYseGlnGlnInaArgAAlaYseGlnGlnInaArg 240
Db 673 AAAAGGAAAGATGAGATATTATAGAGAGATTTATATGCTTTAGAAAATACCACT 732
QY 241 IleGlnLeuProSerGlnAseGlnGlnInaArgAAlaYseGlnGlnInaArg 260
Db 733 ATAGAACTTCATCAGAAAATGAGCGTATATTATATACAGATCATCTTTTACCT 792

Db 853 AGAGATCTATTACAAACAATGTTGACGACGAGGATATATACATTAAGACATCTTGA 912

Qy 301 Gtluylslybapglserileysprogluglulysgluaplyserser 316

Db 913 GAAAGAAAGATGTTCAATTAACCAAGACAAAGAAAGATTAATCT 960

RESULT 7

ARI62974 1482 bp DNA linear PAT 17-OCT-2001

LOCUS ARI62974

DEFINITION Sequence 45 from patent US 6270771.

ACCESSION ARI62974

VERSION ARI62974.1 GI:16233440

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE Unclassified.

AUTHORS 1 (bases 1 to 1482)

TITLE Guerin-Marchand,C. and Druilhé,P.

JOURNAL Peptide sequences specific for the hepatic stages of *P. falciparum* bearing epitopes capable of stimulating the T lymphocytes

FEATURES

Source location/Qualifiers

1..1482

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 2,316-85 Length: 1482

Score: 1546.00 Matches: 313

Percent Similarity: 99.37% Conservative: 1

Best Local Similarity: 99.05% Mismatches: 2

Query Match: 98.91% Indels: 0

DB: 6 Gaps: 0

US-09-837-344-31 (1-316) x ARI62974 (1-1482)

Qy 1 Serapleuglulglnlugargalalysglulysleuglulglnlgnlnserapleu 20

Db 13 ACCGATCTGAACAGACAGACCTGCTTAAGAAAATTCACAGACAAAGCCATTTA 72

Qy 21 Gtluinaapargleualalysglulysleuglulglnlgnlnserapleuglulgln 40

Db 73 GAACAAGATGACTTCTTAAGAAAAGTTACAAAGACACAAAGCATTAGAACAAAG 132

Qy 41 Argargalalysglulysleuglulglnlgnlnserapleuglulglnlugargal 60

Db 133 AGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGACGTGCT 192

Qy 61 Lysglulysleuglulglnlgnlnserapleuglulglnlugargalalysglulys 80

Db 193 AAGAAAAGTTGCAAGACAGACGATTTAGAACAGAGACGTGCTTAAGAAAAG 252

Qy 81 Lysglulglnlgnlnserapleuglulglnaapargleualalysglulysleuglulgln 100

Db 253 TTGCAAGACAAAGCCATTAGAACAAAGATGACTTGAAGAAAAGTTACAAAG 312

Qy 101 Gtlnlnserapleuglulglnlugargalalysglulysleuglulglnlgnlnser 120

Db 313 CAGCAAAAGCATTTTAACAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAAGC 372

Qy 121 Aapleuglulglnlugargalalysglulysleuglulglnlgnlnserapleuglu 140

Db 373 GATTTGAACAGACAGACGTGCTTAAGAAAAGTTGCAAGACAAAGCCATTAGAA 432

Qy 141 Gtlnlugargalalysglulysleuglulglnlgnlnserapleuglulglnlugar 160

Db 433 CAAGACAGACTTCTTAAGAAAAGTTGCAAGAACAAAGCATTAGAACAGAGAG 492

Qy 161 Arggalalysglulysleuglulglnlgnlnserapleuglulglnlugargalalys 180

Db 493 CGTGCTAAAGAAAAGTTGCAAGAACAAAGCATTAGAACAGAGACGTGCTAA 552

Qy 161 Gtluylsleuglulglnlgnlnserapleuglulglnlugargalalysglulysleu 200

Db 553 GAAAGTTGCAAGAACAAAGCCATTAGAACAGAGAGACGTGCTTAAGAAAAGTTG 612

Qy 201 Gtlnlgnlgnlnlugarapleuglulglnlugargalalysglulysleuglulgln 220

Db 613 CAGAGCGCAAGAGATTTAGAACAAAGAGCGCTGACGAAAATAATTTAGAAC 672

Qy 221 Lyslsglulhsgllyaspilleuualagluapleuyrglyargleugluileprola 240

Db 673 AAAAAAGAACATGAGATATATTAGCAGAGATTTATATGTCGTTGAATAACAGCT 732

Qy 241 lileglueuprosersgluasngluargglytyrtyllepnohlglnserseuero 260

Db 733 ATAGAACTTCATCAGAAAAATGAACGTGATATTATATACCAATCAATCTTACCT 792

Qy 261 Gtlnaapamargglyaanserargapserlysgluliserilleglulysrthraen 280

Db 793 CAGCAACAGAGGGAATGAGAGATTCAGAGAAATGCTATATAGAAAACAAAT 852

Qy 281 Arggluserilethrasnvalglulylargargapillehlysglyhlsleuglu 300

Db 853 AGAGATCTATTACAAACAAATGTTGACAGACGACGATATACATTAAGACATCTTGA 912

Qy 301 Gtluylslybapglserileysprogluglulysgluaplyserser 316

Db 913 GAAAGAAAGATGTTCAATTAACCAAGACAAAGAAAGATTAATCT 960

RESULT 8

ARI78263 1482 bp DNA linear PAT 20-APR-2002

LOCUS ARI78263

DEFINITION Sequence 45 from patent US 6319502.

ACCESSION ARI78263

VERSION ARI78263.1 GI:20219401

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE Unclassified.

AUTHORS 1 (bases 1 to 1482)

TITLE Guerin-Marchand,C. and Druilhé,P.

JOURNAL Peptide sequences specific for the hepatic stages of *P. falciparum* bearing epitopes capable of stimulating the T lymphocytes

FEATURES

Source location/Qualifiers

1..1482

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 2,316-85 Length: 1482

Score: 1546.00 Matches: 313

Percent Similarity: 99.37% Conservative: 1

Best Local Similarity: 99.05% Mismatches: 2

Query Match: 98.91% Indels: 0

DB: 6 Gaps: 0

US-09-837-344-31 (1-316) x ARI78263 (1-1482)

Qy 1 Serapleuglulglnlugargalalysglulysleuglulglnlgnlnserapleu 20

Db 13 ACCGATCTGAACAGACAGACGTGCTTAAGAAAAGTTGCAAGAACAAAGCCATTTA 72

Qy 21 Gtluinaapargleualalysglulysleuglulglnlgnlnserapleuglulgln 40

Db 73 GAACAAGATGACTTCTTAAGAAAAGTTACAAAGACCAAGCCATTAGAACAAAG 132

Qy 41 Argargalalysglulysleuglulglnlgnlnserapleuglulglnlugargal 60

Db 133 AGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCATTAGAACAGAGACGTGCT 192

Qy 61 Lysglulysleuglulglnlgnlnserapleuglulglnlugargalalysglulys 80

OY 281 ArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHisIleYsgIlyHisIleuGlu 300
DB 850 AGAAGATCTATTACCAACAAAGTTGAAAGAGCAAGAGGATTTACATTAAGACATCTTGAA 909
OY 301 GluLysIleAspGlySerIleYsPProGluGlnLysGluAspLysSer 316
DB 910 GAAAGAAAGATGCTTCATTAACCAAGACAAAGAAAGATTAATCT 957

RESULT 11
AR162967 1493 bp DNA linear PAT 17-OCT-2001
LOCUS AR162967
Sequence 38 from patent US 6270771.
ACCESSION AR162967
VERSION AR162967.1 GI:16233432
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1493)
AUTHORS Guerin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6270771-A 38 07-AUG-2001,
FEATURES Location/Qualifiers
1..1493
source /organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,896-80 Length: 1493
Score: 1465.00 Matches: 311
Percent Similarity: 98.42% Conservative: 0
Best Local Similarity: 98.42% Mismatches: 5
Query Match: 93.73% Indels: 3
DB: Gaps: 0

US-09-837-344-31 (1-316) x AR162967 (1-1493)

OY 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeu 20
DB 13 AGCATCTTAGAACAAAGACGCTGCTTAAGAAAGTTGCCAAGAACAAAGCGATTTA 72
OY 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlu 40
DB 73 GAACAAATGACCTTGCTTAAGAAAGATTAACAAAGCGAAGCGAATTTAGAACAAAG 132
OY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGluArgAla 60
DB 133 AGACTTGCTTAA-GAAAAGTTGCCAAGAACAAAGCGATCTTAAGAACAAAGACGCT 191
OY 61 LysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAlaLysGlu 80
DB 192 AAAAGAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGACGCTTAAGAAAG 251
OY 81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGlu 100
DB 252 TTGCAGAACAAACAAAGCGATTTAGAACAAATGCTTAAGAAAGAAAGTTCAAGAG 311
OY 101 GlnGlnSerAspLeuGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnSer 120
DB 312 CAGCAAAAGCGATTTAGAACAAAGACGCTTAAGAAAGTTGCCAAGAACAAAGCG 371
OY 121 AspLeuGlnGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlu 140
DB 372 GATTTAGAACAAAGACGCTGCTTA-GAAAAGTTGCCAAGAACAAAGCGATTTAGAA 430
OY 141 GlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArg 160
DB 431 CAAGAGACGCTTGCTTAAGAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGAGAG 490
OY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAlaLys 180

DB 491 CGTCTAAAGAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGACGCTTAA- 549
OY 181 GluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAlaLysGluLysLeu 200
DB 550 GAAAAGTTGCCAAGAACAAAGCGATTTAGAACAAAGACGCTTAAAGAAAGTTG 609
OY 201 GlnGluGlnGlnGlnArgAspLeuGlnGlnGlyArgAlaAspThrLysYsAsnLeuGluArg 220
DB 610 CAAGACCGACGAAAGATTTAGAACAAAGAGCGCTATACGAAAATAATTAGAAAG 669
OY 221 LysLysGluHisIleGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIleProAla 240
DB 670 AAAAGAAACATCGAATATATTAGCAGAGGATTTATATGCTGTTAGAAATACAGCT 729
OY 241 IleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisIleGlnSerSerLeuPro 260
DB 730 ATGAACTTCATCAAGAAATGAAACGCTGATATATATACCATCAATCTTCTTAACCT 789
OY 261 GlnAspAsnArgGluLysSerArgAspSerLysGluIleSerIleIleGluLysThrAsn 280
DB 790 CAGGACAAACGAGAGGATAGTATGAGATTCACAGAAATATCTATTAATAGAAAACAAAT 849
OY 281 ArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHisIleYsgIlyHisIleuGlu 300
DB 850 AGAAGATCTATTACCAACAAATGTTGAAGACGAGAGGATTAATACATAAGGACATCTTGAA 909
OY 301 GluLysIleAspGlySerIleYsPProGluGlnLysGluAspLysSer 316
DB 910 GAAAGAAAGATGCTTCATTAACCAAGAACAAAGAAAGATTAATCT 957

RESULT 12
AR178256 1493 bp DNA linear PAT 20-APR-2002
LOCUS AR178256
Sequence 38 from patent US 6319502.
ACCESSION AR178256
VERSION AR178256.1 GI:20219394
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1493)
AUTHORS Guerin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6319502-A 38 20-NOV-2001,
FEATURES Location/Qualifiers
1..1493
source /organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,896-80 Length: 1493
Score: 1465.00 Matches: 311
Percent Similarity: 98.42% Conservative: 0
Best Local Similarity: 98.42% Mismatches: 5
Query Match: 93.73% Indels: 3
DB: Gaps: 0

US-09-837-344-31 (1-316) x AR178256 (1-1493)

OY 1 SerAspLeuGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeu 20
DB 13 AGCATCTTAGAACAAAGACGCTGCTTAAGAAAGTTGCCAAGAACAAAGCGATTTA 72
OY 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlu 40
DB 73 GAACAAATGACCTTGCTTAAGAAAGATTAACAAAGCGAAGCGAATTTAGAACAAAG 132
OY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgAla 60
DB 133 AGACTTGCTTAA-GAAAAGTTGCCAAGAACAAAGCGATCTTAAGAACAAAGACGCT 191

repeat_region	/rpt_type=tandem 1439.1601 /rpt_type=tandem /rpt_unit="(taaa)n"
repeat_region	1474.1579 /rpt_type=tandem 1479.1588 /rpt_type=tandem complement(1618.1669) /note="AT_rich"
repeat_region	/rpt_type=tandem 1756.2184 /rpt_type=tandem 1775.1954 /rpt_type=tandem /rpt_unit="(taaa)n" complement(1707.2735) /note="AT_rich"
repeat_region	/rpt_type=tandem 2772.2814 /rpt_type=tandem 2790.2865 /rpt_type=tandem 2808.2866 /rpt_type=tandem 3142.3171 /rpt_type=tandem /rpt_unit="(a)n" complement(1274.3347) /note="AT_rich"
repeat_region	/rpt_type=tandem complement(3391.3437) /rpt_type=tandem /rpt_unit="(ta)n" complement(3439.3503) /rpt_type=tandem /rpt_unit="(a)n" 3439.3500 /rpt_type=tandem 3440.3500 /rpt_type=tandem complement(3561.3683) /rpt_type=tandem /rpt_unit="(tataa)n" complement(3710.3745) /note="AT_rich"
repeat_region	/rpt_type=tandem complement(3763.3824) /rpt_type=tandem /rpt_unit="(ta)n" 3763.3811 /rpt_type=tandem complement(3871.4025) /note="AT_rich"
repeat_region	/rpt_type=tandem complement(3922.3987) /rpt_type=tandem /rpt_unit="(ta)n" 3926.3987 /rpt_type=tandem complement(4037.4082) /rpt_type=tandem /rpt_unit="(tataa)n" 4043.4082 /rpt_type=tandem 4043.4080 /rpt_type=tandem complement(4084.4104) /rpt_type=tandem /rpt_unit="(a)n" complement(4135.4181) /rpt_type=tandem /rpt_unit="(ta)n"

	repeat_region	complement(4182..4313)
		/note="AT-rich"
	repeat_region	/rpe_type= tandem
		4364..4466
	repeat_region	/rpe_type= tandem
		4372..4457
	repeat_region	/rpe_type= tandem
		4377..4472
	repeat_region	/rpe_type= tandem
		complement(4512..4655)
		/note="AT-rich"
	repeat_region	/rpe_type= tandem
		complement(4670..4775)
	repeat_region	/rpe_type= tandem
		/rpe_unit=" (ta)n "
		4675..4768
	repeat_region	/rpe_type= tandem
		4870..4912
	repeat_region	/rpe_type= tandem
		complement(4870..4912)
	repeat_region	/rpe_type= tandem
		/rpe_unit=" (ta)n "
		complement(4918..5000)
	gene	
		/locus_tag="PF10_0308"
		join<5162..5350,5488..5535,5617..5802,5902..5981,
		6099..6162,6254..6349,6463..6540,6727..>8973)
	mRNA	
		<5162..>8973
		/locus_tag="PF10_0308"
	CDS	
		join<5162..5350,5488..5535,5617..5802,5902..5981,
		6099..6162,6254..6349,6463..6540,6727..8973)
		/locus_tag="PF10_0308"
		/codon_start=1
		/product="hypothetical protein"

Alignment Scores:

Pred. No.:	1,02e-77	Length:	253001
Score:	1453.50	Matches:	300
Percent Similarity:	91.59%	Conservative:	5
Best Local Similarity:	90.09%	Mismatches:	11
Query Match:	92.99%	Indels:	17
DB:	2	Gaps:	1

US-09-837-344-31 (1-316) x AE014834 (1-253001)

OY	1	SerAapLeuglIngluaArgAlaLysGluYbLeuglInglInserAapleu	20
Db	179807	AGCGATCTAGAACAAGAGACTTAAAGAAAATTACACAGGGGCACAAAGCGATCTAGAACAGAG	179866
OY	21	GIUGIINAPARGLEUALALYSGLUYBLeuglInglInserAapleuglIngu	40
Db	179867	GAACAAAGAGACTTAGTAATAAAGTTACAGGGGCACAAAGCGATCTAGAACAGAG	179928
OY	41	ArgArGaLaLySGLuLYSLeuglInglInglInserAapLeuglIngluaArgAla	60
Db	179927	AGACTAGCTTAAGAAAGATTACAGGGGCACAAAGCGATTTTGACAAAGAGACTTGGT	179986
OY	61	LysGlUyLbLeuglInglInglInserAapleuglIngluaArgAlaLYSGLUYB	80
Db	179987	AAAGAAAGTTCCAGAGCAGCAAAAGCCGATTTAGAACAAAGAGACTTCCTAAAGAAAAAG	180046
OY	81	LeuglInglInglInserAapleuglInglInAPARGLEUALALYSGLUYBLeuglIngu	100
Db	180047	TTCGAAAGAACAAAGCGATTTAACAAAGAGACTTAGTAAGAAAAGTTCCAGAGAA	180106
OY	101	GlInglInserAapleuglInglIngluaArgAlaLYSGLUYBLeuglInglInglInser	120
Db	180107	CAACAAAGCGATTTTGACAAAGAGACTTGTCTAAAGAAAAGTTCCAGAACAAAGAG	180166
OY	121	AapLeuglInglIngluaArgAlaLYSGLUYB-----	131

DB 180167 GATTAGAACAGAGAGACTAGCTAAGAAAAAGTTACAGGGGACAGATGACATTTAGAA 180226
QY 132 -----LeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArg 143
DB 180227 CGAACGAAGCATCTAAAGAAACGTTGCAAGAACAAAGCGATTTAGAACAAAGAGA 180286
QY 144 ArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLys 163
DB 180287 CTTCGTAAGAAAAAGTTGCACAAACAAACGATTTGAAACAAAGAGAGACTGCTAAA 180346
QY 164 GluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLysGluLysLeu 183
DB 180347 GAAAAGTTGCACAAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAAAAGATTG 180406
QY 184 GlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGln 203
DB 180407 CAAGAACAAACAAAGCGATTTAGAACAAAGAGAGAGCTGCTAAAAGATTGCAAGAGCAG 180466
QY 204 GlnAspAspLeuGlnGlnGlnGlnArgAlaAspThrLysLysAspLeuGlnGlnGlnLysGlu 223
DB 180467 CAAGAGATTTGACAAACAAAGCGCTGATACGAAAAAAATTTGAAAGAAAAAGGAA 180526
QY 224 HisGlyAspIleLeuAlaGluAspLeuArgIleArgLeuGlnIleProAlaIleGluLeu 243
DB 180527 CATGGAGATGATTTAGCAGAGATTTATATGGTCGTTTGAATAATCCAGCTATAGAACTT 180586
QY 244 ProSerGluAsnGlnLysGlyTyrTyrIleProHisGlnSerSerLeuProGlnAspAsn 263
DB 180587 CCATCAGAAAAAGAGCGTGATATATATACCAATCATCTCTTTACCTCAGAGACAAAC 180646
QY 264 ArgGlyAsnSerArgAspSerLysGluLysSerIleIleGluLysThrAsnArgLysSer 283
DB 180647 AAGGGCAATATGATGACATTCGAAAGGAAATATCTATATGAAATATCAAAATGAGAAATC 180706
QY 284 IleThrThrAsnValGluLysArgAspIleHisLysGlyHisLeuGlnGlnLysLys 303
DB 180707 ATTACAAACAATGTTGCAAGAGAGAGGATATACATTAAGGAGCATCTTGAAGAAAA 180766
QY 304 AspGlySerIleLysProGlnGlnLysGluLysLysSer 316
DB 180767 GATGTTCAATAAACCAAGAACAAAGAAATTAATCT 180805

RESULT 14
LOCUS ARI62966 954 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 37 from patent US 6270771.
ACCESSION ARI62966
VERSION ARI62966.1 GI:16233431
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 954)
AUTHORS Guerin-Marchand,C. and Druilhé,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6270771-A 37 07-AUG-2001;
FEATURES
source Location/Qualifiers
1..954
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores: 1.35e-38 Length: 954
Pred. No.: 771.50 Matches: 163
Score: 92.59% Conservative: 12
Percent Similarity: 86.24% Mismatches: 13
Best Local Similarity: 49.36% Indels: 1
Query Match: 6 Gaps: 1
DB: 9

US-09-837-344-31 (1-316) x ARI62966 (1-954)

QY 2 AspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGln 21
DB 391 GATGACGATTAATAAATAATATTAAGGCGCAAGACGAAACAAAGCAAGATCTTGA 450
QY 22 GlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnArg 41
DB 451 ---GAAAAGCGCTAAAGAAAGTTACAGGGGCAACAAACGATTCAGAACAAAGAGA 507
QY 42 ArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgArgAlaLys 61
DB 508 CTGCTAAAGAAAAAGTTGCACAAACAAAGCGATTTAGAACAAAGAGAGACTTCTAAA 567
QY 62 GluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLysGluLysLeu 81
DB 568 GAAAAGTTGCACAAACAAAGCGATTTAGAACAAAGAGAGAGCGCTTAAAGAAAAAGTTG 627
QY 82 GlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGln 101
DB 628 CAAGAACAAACAAAGCGATTTAGAACAAAGAGAGAGACTTCTAAAAGAAAGTTGCAAGACAA 687
QY 102 GlnSerAspLeuGlnGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSerAsp 121
DB 688 CAAGCGATTTAGAACAAAGAGAGAGCGCTTAAAGAAAAAGTTGCAAGAACAAAGCGAT 747
QY 122 LeuGlnGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGln 141
DB 748 TTAGAACAAAGAGAGAGCGCTTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAA 807
QY 142 GluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArg 161
DB 808 GAGAGACTTGCTTAAAGAAAAAGTTACAAAGAGAGCAAAACGATTTAGAACAAAGATGACTT 867
QY 162 AlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgArgAlaLysGlu 181
DB 868 GCTTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCGCTTAAAGAA 927
QY 182 LysLeuGlnGlnGlnGlnSerAspLeu 190
DB 928 AGCTTGCAAGAACAAAGCGATTTA 954

RESULT 15
LOCUS ARI178255 954 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 37 from patent US 6319502.
ACCESSION ARI178255
VERSION ARI178255.1 GI:20219393
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 954)
AUTHORS Guerin-Marchand,C. and Druilhé,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes
JOURNAL Patent: US 6319502-A 37 20-NOV-2001;
FEATURES
source Location/Qualifiers
1..954
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores: 1.35e-38 Length: 954
Pred. No.: 771.50 Matches: 163
Score: 92.59% Conservative: 12
Percent Similarity: 86.24% Mismatches: 13
Best Local Similarity: 49.36% Indels: 1
Query Match: 6 Gaps: 1
DB: 9

US-09-837-344-31 (1-316) x ARI178255 (1-954)

QY 2 AspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGln 21
DB 391 GATGACGATTAATAAATAATATTAAGGCGCAAGACGAAACAAAGCAAGATCTTGA 450
QY 22 GlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnArg 41
DB 451 ---GAAAAGCGCTTAAAGAAAGTTACAGGGGCAACAAACGATTCAGAACAAAGAGA 507
QY 42 ArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgArgAlaLys 61
DB 508 CTGCTAAAGAAAAAGTTGCACAAACAAAGCGATTTAGAACAAAGAGAGACTTCTAAA 567
QY 62 GluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgAlaLysGluLysLeu 81
DB 568 GAAAAGTTGCACAAACAAAGCGATTTAGAACAAAGAGAGAGCGCTTAAAGAAAAAGTTG 627
QY 82 GlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGln 101
DB 628 CAAGAACAAACAAAGCGATTTAGAACAAAGAGAGAGACTTCTAAAAGAAAGTTGCAAGACAA 687
QY 102 GlnSerAspLeuGlnGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSerAsp 121
DB 688 CAAGCGATTTAGAACAAAGAGAGAGCGCTTAAAGAAAAAGTTGCAAGAACAAAGCGAT 747
QY 122 LeuGlnGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGln 141
DB 748 TTAGAACAAAGAGAGAGCGCTTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAA 807
QY 142 GluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArg 161
DB 808 GAGAGACTTGCTTAAAGAAAAAGTTACAAAGAGAGCAAAACGATTTAGAACAAAGATGACTT 867
QY 162 AlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnArgArgAlaLysGlu 181
DB 868 GCTTAAAGAAAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGAGAGCGCTTAAAGAA 927
QY 182 LysLeuGlnGlnGlnGlnSerAspLeu 190
DB 928 AGCTTGCAAGAACAAAGCGATTTA 954

```
Db      391 GATGACGATAAAAAAAAAATATATTAAAGCGACAAGAAAAACAGACAAAGATCTTGAA 450
Qy      22  GlnAapArgLeuAlaYsGluYsLeuGlnGlnGlnSerAapLeuGlnGlnArg 41
      : : : : :
Db      451 ---GAAAAAGCAGCTAAAGAAAAGTTACAGGGGCAACAAAGCGATTCAAGAACAAAGAGAGA 507
Qy      42  ArgAlaYsGluYsLeuGlnGlnGlnSerAapLeuGlnGlnGlnArgAlaYs 61
      : : : : :
Db      508 CGTGCTAAAGAAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGCTTGCTAAA 567
Qy      62  GluYsLeuGlnGlnGlnSerAapLeuGlnGlnGlnArgAlaYsGluYsLeu 81
      : : : : :
Db      568 GAAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGCTTGCTAAAAGAAAAGTTG 627
Qy      82  GlnGlnGlnGlnSerAapLeuGlnGlnAapArgLeuAlaYsGluYsLeuGlnGln 101
      : : : : :
Db      628 CAAGAACAAACAAAGCGATTAGAACAAAGAGAGCTTGCTAAAAGAAAAGTTGCAAGAACAA 687
Qy      102 GlnSerAapLeuGlnGlnGlnGlnArgAlaYsGluYsLeuGlnGlnGlnSerAap 121
      : : : : :
Db      688 CAAAGCGATTAGAACAAAGAGAGAGCTTGCTAAAAGAAAAGTTGCAAGAACAAAGCGAT 747
Qy      122 LeuGlnGlnGlnGlnArgAlaYsGluYsLeuGlnGlnGlnGlnSerAapLeuGln 141
      : : : : :
Db      748 TTAGAACAAAGAGAGAGCTTGCTAAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAA 807
Qy      142 GluArgAlaYsGluYsLeuGlnGlnGlnGlnSerAapLeuGlnGlnGlnArg 161
      : : : : :
Db      808 GAGAGACTTGCTAAAAGAAAAGTTACAAAGAGCAGCAAGCGATTAGAACAAAGATAGACTT 867
Qy      162 AlaYsGluYsLeuGlnGlnGlnGlnSerAapLeuGlnGlnGlnGlnArgAlaYsGlu 181
      : : : : :
Db      868 GCTAAAGAAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGAGCTTGCTAAAAGAA 927
Qy      182 LysLeuGlnGlnGlnGlnSerAapLeu 190
      : : : : :
Db      928 AGGTTGCAAGAACAAAGCGATTTA 954
```

Search completed: December 7, 2005, 10:33:18
Job time : 7132.36 sec8

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:22:20 ; Search time 826.854 Seconds

(without alignments)
2547.054 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
Sequence: 1 SDLEQERRAKEKQEQSD.....GHLEKKDGSIKPEQKEDKS 316

Scoring table:

BLASUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Rgapop 6.0 , Rgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 496997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-O=/cgn2.1/USPTO.spool.h/US09837344/runat.05122005.133531.726/app.query.facta_1.718
-DB=N Geneseq -OFMT=fasta -SUPLEX=mg -MTMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNIT8=bits -START=1 -END=1 -MATRIX=blauum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR.SCOR=spc -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344 @CGN 1 1 900 @runat.05122005.133531.726 -NCPU=6 -ICPU=3
-NO MMAP -LARGOUBERY -NEG.SCOR=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -RGAPOP=6
-RGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq_21:*

1: Geneseqn1980s:*\n2: Geneseqn1990s:*\n3: Geneseqn2000s:*\n4: Geneseqn2001s:*\n5: Geneseqn2002s:*\n6: Geneseqn2003s:*\n7: Geneseqn2004s:*\n8: Geneseqn2005s:*\n9: Geneseqn2006s:*\n10: Geneseqn2007s:*\n11: Geneseqn2008s:*\n12: Geneseqn2009s:*\n13: Geneseqn2010s:*\n14: Geneseqn2011s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1549	99.1	950	2	AAQ28115
2	1549	99.1	1496	2	AAQ28119
3	942.5	60.3	1566	2	AAQ05140
4	771.5	49.4	988	2	AAQ28117

5	740.5	47.4	1371	12	ADO21941	Ado21941 LSA-NRC(H
6	740.5	47.4	1374	12	ADO21919	Ado21919 LSA-NRC(H
7	606.5	38.8	1320	2	AAQ08916	Aag08916 Plasmodiu
8	369	25.1	30191	12	ADQ97654	Adq97654 Mouse can
9	369	23.6	3662	4	ABL05275	Ab105275 Drosophi1
10	355	22.7	3012	12	ADN12155	Adn12155 Herpesvir
11	354	22.6	110000	14	AD212814	Continuation (3 of
12	353.5	22.6	3127	12	ADN12154	Adn12154 Kaposi's
13	329.5	21.1	4320	13	ADP12190	Adp12190 Trichohe
14	329.5	21.1	4320	14	ADM00406	Adm00406 Epithelia
15	327.5	21.0	3489	3	AAA30290	Aaa30290 Kaposi's
16	327.5	21.0	3489	6	AAAF82901	Aaf82901 Nucleotid
17	327.5	21.0	3489	6	ABA93487	Aba93487 Kaposi's
18	327.5	21.0	3489	12	ADJ65095	Adj65095 HIV8 DNA
19	327.5	21.0	3489	14	ADG17330	Adg17330 T. gondii
20	327.5	21.0	32207	2	AAV73805	Aav73805 KSHV LUR
21	327.5	21.0	137507	2	AAV19941	Aav19941 KSHV long
22	327.5	21.0	137508	12	ADN12162	Adn12162 Human her
23	326	20.9	1835	2	AAK91242	Aak91242 T. gondii
24	326	20.9	1835	4	AAK42565	Aak42565 T. gondii
25	326	20.9	1835	10	ADG17330	Adg17330 T. gondii
26	321	20.5	1662	4	ABL05275	Ab105275 Drosophi1
27	317.5	20.3	9551	2	AAZ22301	Aaz22301 cDNA enco
28	315.5	20.2	4755	13	ACN37449	Acn37449 Tumour-as
29	311	19.9	2479	11	ADM03281	Adm03281 Human CDN
30	306	19.8	2517	13	ADU01785	Adu01785 Novel hum
31	306	19.6	7078	4	ABL03681	Ab103681 Drosophi1
32	306	19.6	13085	4	ABL03680	Ab103680 Drosophi1
33	303.5	19.4	26214	12	ADQ97246	Adq97246 Mouse can
34	302.5	19.4	1686	2	AAQ87587	Aaq87587 DNA encod
35	301	19.3	945	10	ADF02040	Adf02040 Bacterial
36	301	19.3	3543	10	ADF01979	Adf01979 Bacterial
37	296.5	19.0	8973	12	ADP28653	Adp28653 Human sec
38	292	18.8	1657	13	ADX15088	Adx15088 Plant ful
39	292	18.7	42533	8	ADA56114	Ada56114 Mouse Bat
40	292	18.7	42533	9	ADA02476	Ada02476 Mouse Bat
41	292	18.7	42533	10	ADB72215	Adb72215 Mouse Bat
42	291.5	18.7	1928	11	ADM03216	Adm03216 Human CDN
43	290.5	18.6	6617	12	ADL12539	Adl12539 Human ste
44	289	18.5	210	1	AAH80899	Aah80899 Sequence
45	289	18.5	2772	5	AAK90525	Aak90525 DNA encod

ALIGNMENTS

RESULT 1
AAQ28115
ID AAQ28115 standard; DNA; 950 BP.

XX	AC	AAQ28115;				
XX	DT	25-MAR-2003 (revised)				
XX	DT	08-FEB-1993 (first entry)				
XX	DE	P. falciparum LSA-R-NR coding sequence.				
XX	KM	Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;				
XX	KW	paludism; liver stage-specific antigen; ss.				
XX	OS	Plasmodium falciparum.				
XX	FT	Key	Location/Qualifiers			
XX	FT	repeat_region	3..629			
XX	FT	/*cag= a				
XX	FT	/rpt_type= TANDEM				
XX	FT	repeat_unit	3..53			
XX	FT	/*cag= b				
XX	FN	W09213884-A1.				
XX	PD	20-AUG-1992.				
XX	PF	05-FEB-1992; 92WO-FR000104.				

XX 05-FEB-1991: 91FR-00001286.
 XX (INSP) INST PASTEUR.
 XX Guerimarchand C, Druilhe P;
 XX MPI. 1992-299985/36.
 DR P-PSDB; AAR26941.
 XX
 PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
 PT vaccination against, treatment of and diagnosis of malaria.
 XX
 PS Claim 17; Fig 2; 81pp: French.
 XX
 CC A genomic DNA bank of P. falciparum EcoRI fragments prepared in lambda
 CC gII1 was used to transform E.coli. The expression library was screened
 CC with human antisera against antigens of all stages of P. falciparum. The
 CC library was reselected with antibodies affinity- purified on a clone
 CC which was able to recognise antibodies specific to the hepatic phase.
 CC About 40 clones were detected which produced a characteristic LSA
 CC epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
 CC containing a 12-repeat region followed by a non-repeat region. Preferred
 CC antigenic polypeptides of the invention are derived from the amino acid
 CC sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 950 BP; 456 A; 127 C; 206 G; 161 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,06e-124 Length: 950
 Score: 1549.00 Matches: 314
 Percent Similarity: 99.378 Conservative: 0
 Best Local Similarity: 99.378 Mismatches: 2
 Query Match: 99.108 Indels: 0
 DB: 2 Gaps: 0

US-09-837-344-31 (1-316) x AAQ28115 (1-950)

QY 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnInserAspLeu 20
 DB 3 ACCGATCTAGAAACAAGAGAGACGCTGTAAGAAAGTTCCAAAGAACAAAGCCGATTTA 62
 QY 21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnInserAspLeuGlnGln 40
 DB 63 GAACAAAGATAGACTTCTTAAGAAAGTTTCAAGAGCACAAAGCATTTAGAACAAAG 122
 QY 41 ArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlnGluArgAla 60
 DB 123 AGACTTGCTAAAGAAAGTTGCAAGAAACAAGCGATCTAGAACAAAGAGACGCTGCT 182
 QY 61 LysGluLysLeuGlnGlnInserAspLeuGlnGlnGluArgAlaLysGluLys 80
 DB 183 AAAGAAAGTTGCAAGAAACAAGCGATTTAGAAACAAGAGACGCTGTAAGAAAG 242
 QY 81 LeuGlnGlnGlnInserAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGln 100
 DB 243 TTGCAAGAAACAAGAAAGCGATTTAGAACAAAGATGACTTCTTAAGAAAGTTTCAAGAG 302
 QY 101 GlnGlnInserAspLeuGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnInser 120
 DB 303 CAGCAAAAGCGATTTAGAACAAAGAGACGCTGTAAGAAAGTTGCAAGAAACAAGAAC 362
 QY 121 AspLeuGlnGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGln 140
 DB 363 GATTTAGAAACAAGAGACGCTGTAAGAAAGTTTCAAGAAACAAGCGCGATTAGAA 422
 QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnGln 160
 DB 423 CAAGAGAGACTTCTTAAGAAAGTTGCAAGAAACAAGCATTTAGAACAAAGAGAA 482
 QY 161 ArgAlaLysGluLysLeuGlnGlnGlnInserAspLeuGlnGlnGluArgArgAlaLys 180
 DB 483 CGTGCTAAAGAAAGTTGCAAGAAACAAGCGATTTAGAACAAAGAGACGCTGTAAG 542

QY 181 GluLysLeuGlnGlnGlnInserAspLeuGlnGlnGluArgArgAlaLysGluLysLeu 200
 DB 543 GAAGATTGCAAGAAACAAGCGATTTAGAACAAAGAGACGCTGTAAGAAAGTTG 602
 QY 201 GlnGlnGlnGlnInserAspLeuGlnGlnGluArgArgAlaLysGluLysLeuGlnGln 220
 DB 603 CAAGAGCAAGAAAGATTTAGAACAAAGAGCGCTGATCGAAAAAATTTAGAAAGA 662
 QY 221 LysLysGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
 DB 663 AAAAAGGAACATGAGATATTTAGCAAGAGATTTATATGCTGTTAGAAATACAGCT 722
 QY 241 ILeGluLeuProSerGlnAsnGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
 DB 723 ATGAGACTTCCATGCAAGAAATGAAAGTGGATATTTATATCCATCAATCTTCTTACCT 782
 QY 261 GlnAspAsnArgGlnLysSerArgAspSerLysGlnGlnGlnGlnGlnGlnGlnGln 280
 DB 783 CAGGCAACAGAGGGAATAGTAGAGATTTCCAAAGAAATATCTATATGAAACAAAT 842
 QY 281 ArgGlnSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 300
 DB 843 AGAGATCTATTTACAAAGAAAGTTGAAAGAGCAAGGATATACATTAAGAGCATCTTGA 902
 QY 301 GluLysLysAspGlnSerLysLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
 DB 903 GAAAGAAAGATGCTTCAATTAACCAAGAACAAAGAAAGATTAATCT 950

RESULT 2

AAQ28119 standard; DNA; 1496 BP.

AAQ28119;

25-MAR-2003 (revised)

08-FEB-1993 (first entry)

P. falciparum LSA gene 3' region.

Malaria; hepatocyte, sporozoite; T-cell epitope; B-cell epitope;

paludism; liver stage-specific antigen; ss.

Plasmodium falciparum.

Key Location/Qualifiers

CDS 1..1482

FT /*tag= C

FT /product= "LSA_C-terminal_region"

FT repeat_region 37..639

FT /*tag= a

FT repeat_unit 37..87

FT /*tag= b

FT /*type= TANDEM

MO9213884-A1.

20-AUG-1992.

05-FEB-1992; 92WO-FR000104.

05-FEB-1991; 91FR-00001286.

(INSP) INST PASTEUR.

Guerimarchand C, Druilhe P;

MPI: 1992-299985/36.

P-PSDB; AAR26941.

Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
 PT vaccination against, treatment of and diagnosis of malaria.

|||||
Db 720 CAAGGCGATTTAGAACAGAGACGCTGCTTAAGAAAAGTTGCAAGAACAAAGCGAT 779
Qy 122 LeuGIuGIuGIuAArgAlaIyGluIyLeuGIuGIuGIuInserAspLeuGIuGIu 141
Db 780 TTAGAACAGAGACGCTGCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAA 839
Qy 142 GIuAArgAlaIyGluIyLeuGIuGIuGIuInserAspLeuGIuGIuGIuAArg 161
Db 840 GAGGACTTGTCTTAAGAAAAGTTACAGAGCAGCAAGCGATTTAGAACAGATGACTT 899
Qy 162 AlAlyGluIyLeuGIuGIuGIuInserAspLeuGIuGIuGIuAArgAlaIyGlu 181
Db 900 GCTTAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCTTAAGAAA 959
Qy 182 LyLeuGIuGIuGIuInserAspLeu 190
Db 960 AGCTTGCAAGAACAAAGCGATTTA 986
RESULT 5
ADO21941
ID ADO21941 standard; DNA; 1371 BP.
XX
AC ADO21941;
XX
DT 12-AUG-2004 (first entry)
XX
XX LSA-NRC(H) construct DNA derived from Malaria parasite LSA-1.
DE
XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KW parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KW malaria parasite P. falciparum; LSA-NRC(H) construct; de; gene.
XX
OS Plasmodium falciparum.
OS Synthetic.
XX
FH Location/Qualifiers
FT 1..1371
FT /product= "LSA-NRC(H) construct protein derived from
FT /tag= a
FT /product= "Malaria parasite LSA-1"
XX
XX WO200404167-A2.
XX
XX 27-MAY-2004.
XX
XX 12-NOV-2003; 2003WO-US036011.
XX
XX 12-NOV-2002; 2002US-0425719P.
XX
XX (REED-) REED ARMY INST RES WALTER.
XX
XX Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
XX Barbosa A;
XX WPI; 2004-420309/39.
XX
XX P-PSDB; ADO21942.
XX
XX Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
XX agent, preferably for diagnosing malaria, comprises a liver stage antigen
XX (LSA-1) epitope.
XX
XX
XX Claim 16; SEQ ID NO 25; 90pp; English.
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
XX comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
XX protein is found within the parasitophorous vacuole (PV) of Plasmodium
XX falciparum. The polypeptide of the invention demonstrates antimalarial
XX activities and may be useful as a vaccine during immunotherapy and as a
XX diagnostic agent, preferably for diagnosing malaria. The current sequence
XX is that of the LSA-NRC(H)Mnt construct DNA of the invention. The
XX construct is harmonised for expression in Escherichia coli and comprises
XX the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria

CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
XX
XX Sequence 1371 BP; 539 A; 358 C; 283 G; 191 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1,18e-54 Length: 1371
Score: 740.50 Matches: 177
Percent Similarity: 67.34% Conservative: 23
Best Local Similarity: 59.60% Mismatches: 66
Query Match: 47.38% Indels: 32
DB: 12 Gaps: 7
US-09-837-344-31 (1-316) x ADO21941 (1-1371)
Qy 21 GIuGIuAspAArgLeuAlaIyGluIyLeuGIuGIuGIuInserAspLeuGIuGIu 40
Db 16 GAAAAGACGCAAAATTTACAAAGCAATCTCCGCGCAGCTTCAAC-----AGCGGC 69
Qy 41 AArgAArgAlaIyGluIyLeuGIuGIuGIuInserAspLeuGIuGIuGIuAArgAla 60
Db 70 AACCGCATCAACAGGAAAAGCATGAGAAACATGTGCTGAGCCCAACTCTACGAG 129
Qy 61 LyGluIyLeuGIuGIuGIuInserAspLeuGIuGIuGIuAArgAlaIyGluIy 80
Db 130 AAGACTTAAAA-CACGAAAACAAATTTGACCAAGCAAGAAAGCTGACAT--- 185
Qy 81 LeuGIuGIuGIuInserAspLeuGIuGIuAspAArgLeuAlaIyGluIyLeuGIuGIu 100
Db 186 -----GAGCAGAGTTTAAACGATATCCAGACCACTTTAAATCCCTCT 230
Qy 101 GIuGIuInserAspLeuGIuGIuAArgAlaIyGluIyLeuGIuGIuGIuInser 120
Db 231 -----GCGCAACCTCGCGCTTCCGAGACATCTT---TCTCAAGAAA 272
Qy 121 AspLeuGIuGIuGIuAArgAlaIyGluIyLeuGIuGIuGIuInserAspLeuGIu 140
Db 273 CAAGCTAACAGAAAGGCAACATGTAATCATCATCAACGACGACGATGAC----- 327
Qy 140 uGIuGIuAArgAlaIyGluIyLeuGIuGIuGIuInserAspLeuGIuGIuGIuA 160
Db 328 -AAAAAAATACATTAAAGCCAGATGAAATCCGACAGAAAGACTTCAAA----- 378
Qy 160 GArgAlaIyGluIyLeuGIuGIuGIuInserAspLeuGIuGIuGIuAArgAlaIy 180
Db 379 -----GAAAACCTGCTGACAGCAGTGGACTGGACCTGAAACGAGCGCTCGCTAA 428
Qy 180 sGIuIyLeuGIuGIuGIuInserAspLeuGIuGIuGIuAArgAlaIyGluIy 200
Db 429 AGAAAAGCTC-----CAGAGCGGCTCGCTTAAGAAAAGCT 464
Qy 200 uGIuGIuGIuAArgAspLeuGIuGIuAArgAlaIyGluIyLeuGIuGIuGIuA 220
Db 465 CAGAGACCAACAGCGGCACTGGAAACGCCAAGGCTGACACGAAAAAAACCTGGAAGC 524
Qy 220 gIyLeuGIuGIuGIuAspLeuIleuAlaIyAspLeuTyGlyAArgLeuIleProAl 240
Db 525 CAAAAAGAACACGCGCACTCTGCGTGAAGACCTTACGCGCGCTGAAAATCCAGC 584
Qy 240 AlLeuGIuLeuProSerGIuAsnGIuAArgIyTyTyTleProHisInserSerLeuPr 260
Db 585 TATCGAATCTCCATCCGAAAACGAAACCGGCTACTACATCCACACAGAGCGCTGCC 644
Qy 260 OGInAspAsnAArgGlyAsnSerAArgAspSerLyGluIleSerIleIleGluTyThrAs 280
Db 645 ACAAGATTAATCGCGGAATCTCCCGCAGATGAAGAAATCAGCATCTCAAAAAACCAA 704
Qy 280 nArgIuSerIleThrThrAsnValGIuIyAArgAArgAlaIyGluIyGluIy 300
Db 705 CCGGAAAGCAATTAACCAACGAGGCGCGGACATCCACAAAGGCACTCGA 764
Qy 300 uGIuIyLeuAspGIySerIleTyPProGIuGIuGIuAArgAspSer 316
Db 765 AGAAAAGAAAGCGCTCCATCAACCAAGAAAGAAAGCAAAAGC 813


```
XX OS Plasmodium falciparum; (pLSARPLS.14L.1).
XX PN WO9428930-A1.
XX PD 22-DEC-1994.
XX PF 10-JUN-1994; 94WO-US006652.
XX PR 11-JUN-1993; 93US-00075783.
XX PR 09-JUN-1994; 94US-00257073.
XX PA (VIRO-) VIROGENETICS CORP.
XX PI Paolucci E, De Talsone C, Tine JA,
XX DR WPI, 1995-036113/05.
XX PT Recombinant poxvirus confg. Plasmodium DNA in non-essential region -
XX PT useful in vaccines against malaria and for prodn. of Plasmodium
XX PT immunogens.
XX PS Claim 3; Fig 11; 183pp; English.
XX CC AA080916 is the P. falciparum LSA-1 repeatless gene cDNA sequence. New
XX CC recombinant poxviruses containing either the SERA, ABRA, P1nsp10, AMA-1,
XX CC P1a25, P1a16, CSP, P1fSP2, LSA-1, LSA-1 repeatless, MSA-1, MSA-1 (N-
XX CC terminal p83 or C-terminal p942) genes, or a combination of these in non-
XX CC essential regions of their genomes are claimed. These poxviruses (pref.
XX CC with a virulence reducing genomic deletion or disruption) can be used as
XX CC vaccines against malaria and for the prodn. of Plasmodium immunogens.
XX CC These viruses provide multicomponent, multistage vaccines due to their
XX CC expression of sporozite, liver stage, blood stage and sexual stage
XX CC proteins. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-
XX CC OCT-2003 to standardize OS field)
SQ Sequence 1320 BP; 619 A; 128 C; 227 G; 346 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.88e-43 Length: 1320
Score: 606.50 Matches: 131
Percent Similarity: 63.60% Conservative: 28
Best Local Similarity: 52.40% Mismatches: 56
Query Match: 38.80% Indels: 35
DB: Gaps: 4
US-09-837-344-31 (1-316) x AA080916 (1-1320)
QY 89 G1UC1nAaPArgLeuAlaLybG1uLyLeuG1nG1nG1nSerAaPLeuG1nG1n 108
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 88 GAAAAAATGATAAAATCAATTAATCTTAAGTGGTCTTCTTAAT-----TCTAGG 141
109 ArgArgAlaLybG1uLyLeuG1nG1nG1nSerAaPLeuG1nG1nG1nG1nG1n 128
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 142 AATGAAATTAATGAGAAAGCAGAAAGAAACAGCTTTATCTCAATTCATAGAC 201
129 LybG1uLyLeuG1nG1nG1nG1nSerAaPLeuG1nG1nG1nG1nG1nG1n 144
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 202 AAAAATTAATTAATGAATTAATTAATTTTCATTAAGGTAAGATGATGCT 261
145 -----AlaLyb 146
262 AATGTAATAATGTGTACAAATTTCAAAAGCTTTTAAAGAACTTGATGTTTCA 321
QY G1uLyLeuG1nG1nG1nG1nSerAaPLeuG1nG1nG1nG1nG1nG1nG1n 166
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 322 GAGAAATATCTCTTAAGAAATTAATTAATTAAGGAAATTAATTAAGACACATA 381
167 G1nG1uG1nG1nSerAaPLeuG1nG1nG1nG1nG1nG1nG1nG1nG1n 186
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 382 ATAAATGATGATGATGAT-----AAAAAAATATATTAAAGGCAAGCAAGAAACAGA 435
187 G1nSerAaPLeuG1nG1nG1nG1nG1nG1nG1nG1nG1nG1nG1nG1nG1n 206
```

```
DB 436 CAAGAAATCTTCAAGAAAAAGCCGCCATCTAA----- 471
QY 207 LeuG1uG1nArgLybAlaAaPThLybAaPLeuG1uArgLybG1uH1eG1yAaP 226
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 472 ---GAAACGAGAAAGCTGTACGAAAAAAATTTGAAAAAGAAAGAAATGGAGAT 528
QY 227 IleLeuAlaG1uAaPLeuThyArgLeuG1uLeProAla11eG1uLeuProSerG1u 246
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 529 GTATTAGCAGAGATTTATATGTCGTTTGAATACACAGCTATAGAACTTCATCAGAA 588
247 AaG1uArgLybThyTyr11eProH1eG1nSerSerLeuProG1nAaPLeuArgLyb 266
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 589 AATGAACGTGATATATATATACATCATCTCTTTACTGAGCAACAGAGGAAT 648
QY 267 SerArgAaPLeuSerLybG1uLeSer11eG1uLybThaAaPLeuG1uSer11eThr 286
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 649 AGTAGAGATTCAGAGAAATATCTATATAGAAAAAACAATAGAAATCTATTACAA 708
QY 287 AaThyAlG1uG1nArgArgAaP11eH1eLybG1uH1eLeuG1uLybAaPLeuG1ySer 306
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 709 AATGTTGAAGACGAAAGGATATACATTAAGACATCTTGAAAGAAAGAAAGATGTTCA 768
QY 307 IleLybProG1uG1nLybG1uAaPLeuSer 316
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 769 ATAAACAGAACAAAGAAAGATTAATCT 798
RESULT 8
ADQ97654/c
ID ADQ97654 standard; DNA; 30191 BP.
XX AC ADQ97654;
XX DT 07-OCT-2004 (first entry)
XX DE Mouse cancer associated sequence MD10-022, SEQ ID 631.
XX KM Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX OS Mus musculus.
XX PN WO2004060304-A2.
XX PD 22-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US041389.
XX PR 27-DEC-2002; 2002US-00330773.
XX PA (SAGR-) SAGRES DISCOVERY INC.
XX PI Morris DW, Malandro MS;
XX DR WPI; 2004-543781/52.
XX XX New isolated cancer associated nucleic acids comprising at least 10
XX PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX PT cancers such as leukemia and lymphoma.
XX PS Claim 1; SEQ ID NO 631; 199pp; English.
XX CC The present invention relates to cancer associated sequences (ADQ97025-
XX CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 30191 BP; 7666 A; 7033 C; 6989 G; 8483 T; 0 U; 20 Other;
Alignment Scores:
Pred. No.: 2.52e-23 Length: 30191
Score: 393.00 Matches: 86
```


[illegible]

```
RESULT 11  
ADZ12814_2  
Continuation (3 of 5) of ADZ12814 from base 200001 (Murine cancer-associated genomic DNA)  
WP Sequence split into 5 fragments LOCUS ADZ12814 Accession Adz12814  
WP Fragment Name Begin End  
WP ADZ12814_0 1 11000  
WP ADZ12814_1 100001 21000  
WP ADZ12814_2 200001 31000  
WP ADZ12814_3 300001 41000  
WP ADZ12814_4 400001 417026
```

	Alignment Scores:	Pred. No.:	Length:
Score:	2.25e-19		110000
Percent Similarity:	354.00		Matches: 81
Best Local Similarity:	56.73%		Conservative: 96
Query Match:	25.96%		Mismatched: 113
DB:	22.65%		Indels: 23
Gaps:	14		Gaps: 4

US-09-837-344-31 (1-316) X ADZ12814_2 (1-110000)

[illegible][illegible]

Percent Similarity:	53.72%	Conservative:	87
Best Local Similarity:	29.75%	Mismatches:	103
Query Match:	21.08%	Indels:	65
GB:	13	Gaps:	16

US-09-837-344-31 (1-316) X ADR12190 (1-4320)

```

Oy 2 AspleuengluglIngluArgArglAlaybgluIlys-----LeuengluglIn 16
Db 1804 GAGCTGGGAGCAGAGAGAGGAGGCGGTTCGACGAGCGCGCACCGTCAGATCTCTTAGAGAGAGAG 1863
Oy 17 GlInserApleuengluglInAserAryLeuAllys-----gluIyLeuengluglIn 33
Db 1864 CAGTTTCACGAGAGACCAACCGGAGACGAGAAACGAGATTCAGACTTCCAGAGAGGAA 1923
Oy 34 GlIn-----SerApleuIn 38
Db 1924 GACACAGCTCCAGGGAGAAATCGAAGAAACGGCAACGAGAGAGAGCGCAAGTTCCTTAG 1983
Oy 39 GlngluIATArglAlaybgluIyLeuengluglInglInserApleuengluglInIyArg 58
Db 1984 GAGCAAGAGCGACGCTCGGACAGAAACGGGAAAGACGAGCGCGCTCAAGAACAGAGAGGA 2043
Oy 59 ArglAlaybgluIyS-----LeuengluglInglInserApleuengluglInIyArg 75
Db 2044 GAATTCCTCAAGAGAGAGAGAGAGACCTCCACAGAAACCGCAAGAAAGAACTTCGACAGAAATGC 2103
Oy 76 ArglAlaybgluIyLeuengluglInglInserApleuengluglInAserAryLeuAllys 95
Db 2104 GACAGAAAATCTCGGAACAAAGAGGCGCGCACGACCGTGAAGAAAGACGCTTAGGGCTT 2163
Oy 96 gluIySLeuengluglInglInserApleuengluglInIyArg-----ArgAlayS 112
Db 2164 CAGAGG---CGGACCAAGAGATTCGTCGGGAAACAAGAACGCCACCTGAACGTGAGGAA 2220
Oy 113 gluIySLeuengluglInglInserApleuengluglInIyArg-----ArgAlayS 129
Db 2221 GAGCGAGCTCGCGGACAGACCAATCCCGCGGAAACAAGAACGCCACACAGAAACGTGAGGAA 2280
Oy 130 gluIySLeuengluglInglInserApleuengluglInIyArg-----ArgAlayS 146
Db 2281 GAGCGAGCTCGCGGACAGACCAATCCCGCGGAAACAAGAACGCCACACAGAAACGTGAGGAA 2340
Oy 147 gluIySLeuengluglInglInserApleuengluglInIyArg-----ArgAlayS 163
Db 2341 GAGCGAGCTCGCGGACAGACCAATCCCGCGGAAACAAGAACGCCACACAGAAACGTGAGGAA 2400
Oy 164 gluIySLeuengluglInglInserApleuengluglInIyArg-----ArgAlayS 180
Db 2401 GAGCGAGCTCGCGGACAGACCAATCCCGCGGAAACAAGAACGCCCGCTGAGAGCTGAGGAA 2460
Oy 181 gluIySLeuengluglInglInserApleuengluglInIyArgAlaybgluIyLeu 200
Db 2461 GAGCGAGCTCGCGGACAGACCAATCCCGCGGAAACAAGAACGCCAC---CAGGAAACGTGAG 2511
Oy 201 GlngluIuglInIyArgApleuengluglInIyArgAlaybThlyLybAanleuInIyArg 220
Db 2518 GAGAGCGAGCTCGCGGACAGACCAATCCCGCGG---GAACAGAAACGCCCTGAGAGCT 2574
Oy 221 LuylaybgluIhAglYAvrIleuAlagIuApleuIyArgIyLeuengluglInIyProAla 240
Db 2575 GAGGAAAGAGCAG-----CTCGCGGACCAATCATTCCTCGCGGAGAGCA----- 2611
Oy 241 IlegIuLeuProSerGluAangluIyArgIyTyTyTleerohIglInserSerLeuPro 260
Db 2617 ---GAGCTCAGACCGGACAGAAATTCATCGTAGAGAAAGACCGCGCACAGAGACCTGAG 2673
Oy 261 GlnArapAanATxGlyAanSerAryAserIyvguIleSerIleIlegIuYerThyAn 280
Db 2674 GAAGAGCAGCGTGGCCAGAGCGGAGACCGTTGAGGGTGAAGAGACGTTTGGCGAGCAG 2733
Oy 281 ArgIu-----SerIleThrThraValGlu 289

```

Db	2734	CGAAGAGGAAGCGACGCCGCCCGCCAGGAATGTGACAGAAATTTACACGGGAATCAGG	2739		
Qy	290	GLYRGTGAAPRIENHSLYGLYHISLEUGLUGLULYLYSAPRGLYSERILEYSPRO	309		
Db	2794	GTCCGCCAGGAACGTGAGGAAGACGCGCTGGCGGACAGAAAG-----CTCCGCAGG	2844		
Qy	310	GLUGLULYLYS	312		
Db	2845	GAACAAAGG	2853		
RESULT 14					
ADW00406					
ID	ADW00406	standard; CDNA; 4320 BP.			
XX					
XX	ADW00406;				
XX					
DT	07-APR-2005	(first entry)			
XX					
DE	Epithelial hair follicle related antigen coding sequence, SEQ ID 159.				
XX					
KW	Endocrine-Gen.; Cardiovascular-Gen.; Vuninary; Osteopathic; Cardiant;				
KW	Gene Therapy; antibody production; antibody therapy; antigen;				
KW	hair growth; arteriosclerosis; osteoporosis; bone injury; angina; burns;				
KW	wound healing; gene; ss.				
XX					
OS	Mus musculus.				
XX					
PH	Key	Location/Qualifiers			
FT	CDS	1..4320			
FT	/*tag= a				
FT	/product= "Epithelial hair follicle related antigen"				
XX					
PN	US2005008602-A1.				
XX					
PD	13-JAN-2005.				
XX					
PF	07-JAN-2004; 2004US-00754079.				
XX					
PR	05-JUN-2000; 2000JP-00166903.				
PR	04-JUN-2001; 2001WO-JP004691.				
PR	13-NOV-2001; 2001JP-00347338.				
PR	13-NOV-2001; 2001JP-00347340.				
PR	05-DEC-2001; 2001JP-00371175.				
PR	05-DEC-2001; 2001JP-00371366.				
PR	23-MAY-2002; 2002JP-00155922.				
PR	08-JAN-2003; 2003JP-00001891.				
XX					
PA	(SUME) SUMITOMO ELECTRIC IND LTD.				
XX					
PI	Hirai Y, Takebe K;				
XX					
DR	WPI: 2005-080487/09.				
DR	P-PSDB; ADW00404.				
XX					
PT	New antibody specifically recognising an antigen of 220kda present in				
PT	epithelial new follicles, useful for evaluating and inducing hair growth				
PT	promoting activity and morphogenesis.				
XX					
PS	Example 19; SEQ ID NO 159; 87p; English.				
XX					
CC	The present invention relates to a novel antibody, or its fragment, which				
CC	specifically recognizes an antigen of about 220KDa (ADW00404) present in				
CC	new epithelial hair follicles. The antigen is specifically expressed				
CC	during the growth period of an imago or the developing period of a fetus.				
CC	Also claimed is a method for the evaluation of hair growth promoting				
CC	activity, comprising incubating skin tissue derived from a mammal in the				
CC	presence of a substance to be tested to promote hair growth, recovering				
CC	the skin tissue, reacting the skin tissue with the antibody cited above,				
CC	and detecting the antibody, or its fragment, that reacts with the skin				
CC	tissue. The method and antibody of the invention are useful for				
CC	evaluating and inducing hair growth promoting activity, treating and				
CC	ameliorating symptoms of diseases associated with abnormal morphogenesis,				
CC	inducing cardiovascular revascularization, regeneration and endothelial				

Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;

US-09-837-344-31 (1-316) x AAA30290 (1-3489)

```

Db      2473 TTAGAGGAGCAGAGCAGAGCTTAGAGAGCAGAGCAGAGCTTAGAG-----GAG 2523
QY      220  ArgIySlySgIuH1sgIyAepIleuA1agIuApeIuTyrgIyArSleuGluIlePro 239
      ::::: ||| :::: |||:::
Db      2524 CAGGAGGTGGAAGAGCAGAGCAGAGGTGGAAGAGCAGAGCAGAGGAGAGCAG 2583
QY      240  AlaillegIuleuProSerGIuAengIuArgGIyTyrtIleProH1sgIuSerIleu 259
      ::::: ||| :::: |||:::
Db      2584 GAATTAGAGGAGGTGAGGAGCAGAGCAGAGCAGAGGAGCAGAGGAGCAGAGCTTA 2643
QY      260  ProGIuApeAenArGgIyAenSerArGApsSerIySgIuIleSerIleIlgIuIyThr 279
      ::::: ||| :::: |||:::
Db      2644 GAGGAGGTGGAAGAGCAGAGCAGAGCAGAGCTTAGAGAGGTGGAAGAGCAGAGAGCAG 2703
QY      280  AsnArGSIuSerIleThrThrAsnValGIuGIyArGAArgApeIleH1slySgIyH1leu 299
      ||| ||| ||| |||:::
Db      2704 GAGTTAGG-----GAGGTGGAAGAGCAGAGCAGAGCAGAGCTTAGAGAGGTG 2751
QY      300  GIuGIuIySlyApsGlySerIleIyApeProGIuGIuIuIySgIu 313
      ||| ||| ::::: |||:::
Db      2752 GAAGAGCAGAGCAGAGCAGAGGTGGAACAGCAGAGCAGAGCAGAG 2793

Search completed: December 7, 2005, 07:51:15
Job time : 888.854 secs

```

Search completed: December 7, 2005, 07:51:15
Job time : 888.854 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:30:56; Search time 251.048 Seconds

(without alignments)
2237.454 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
Sequence: 1 SDEGERRAKEKLOEQOSD.....GHLEKKDGSIKPEQKEDKS 316

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Rgapop 6.0 , Rgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.apool/h/US09837344/runat_05122005_133533_763/app.query.fasta_1.718
-DB=Issued_Patents_NA -OPMT=fastgap -SUFFIX=ent -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09837344 @CGN 1.1 237 @runat_05122005_133533_763 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBJECT -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the chance being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	99.1	950	3	US-08-098-327E-32
2	1549	99.1	950	3	US-08-462-625-32
3	1549	99.1	1482	3	US-08-098-327E-41
4	1549	99.1	1482	3	US-08-462-625-41
5	1546	98.9	1482	3	US-08-098-327E-45
6	1546	98.9	1482	3	US-08-462-625-45
7	1465	93.7	1493	3	US-08-098-327E-38
8	1465	93.7	1493	3	US-08-462-625-38
9	771.5	49.4	954	3	US-08-098-327E-37

10	771.5	49.4	954	3	US-08-462-625-37	Sequence 37, App1
11	771.5	49.4	988	3	US-08-098-327E-34	Sequence 34, App1
12	771.5	49.4	988	3	US-08-462-625-34	Sequence 34, App1
13	606.5	38.8	1320	2	US-08-257-073-15	Sequence 15, App1
14	344	22.0	1827	3	US-09-270-767-1308	Sequence 1308, App
15	344	22.0	1827	3	US-09-270-767-16590	Sequence 16590, A
16	327.5	21.0	3489	2	US-08-728-323A-1	Sequence 1, App11
17	327.5	21.0	3489	3	US-09-298-568-1	Sequence 1, App11
18	327.5	21.0	3489	3	US-09-410-399-1	Sequence 1, App11
19	327.5	21.0	3489	3	US-09-834-273-1	Sequence 1, App11
20	327.5	21.0	32207	2	US-08-770-379-20	Sequence 20, App1
21	327.5	21.0	32207	3	US-08-757-669A-20	Sequence 20, App1
22	327.5	21.0	32207	3	US-09-230-371A-20	Sequence 20, App1
23	326	20.9	1835	3	US-09-216-393B-80	Sequence 80, App1
24	319	20.4	767677	3	US-09-949-016-12347	Sequence 12147, A
25	319	20.4	767677	3	US-09-949-016-17361	Sequence 17361, A
26	317.5	20.3	9551	2	US-08-056-200-93	Sequence 93, App1
27	317.5	20.3	9551	2	US-08-800-644-93	Sequence 93, App1
28	301	19.3	945	3	US-09-543-681A-2325	Sequence 2325, App
29	301	19.3	3543	3	US-09-543-681A-2264	Sequence 2264, App
30	296.5	19.0	29927	3	US-09-949-016-11814	Sequence 17474, A
31	296.5	19.0	29927	3	US-09-949-016-17474	Sequence 17474, A
32	296.5	19.0	29927	3	US-09-949-016-17475	Sequence 17475, A
33	296.5	19.0	37802	3	US-09-949-016-12639	Sequence 12639, A
34	290.5	18.6	6617	3	US-09-976-594-268	Sequence 268, App
35	288	18.4	2093	3	US-10-104-047-1665	Sequence 1665, App
36	286.5	18.3	1522	3	US-10-104-047-1093	Sequence 1093, App
37	284	18.2	7463	3	US-09-902-540-928	Sequence 928, App
38	281.5	18.0	1965	3	US-09-902-540-1894	Sequence 1894, App
39	278.5	17.8	29927	3	US-09-949-016-11814	Sequence 1814, A
40	278.5	17.8	29927	3	US-09-949-016-17474	Sequence 17474, A
41	278.5	17.8	29927	3	US-09-949-016-17475	Sequence 17475, A
42	278.5	17.8	37802	3	US-09-949-016-12639	Sequence 12639, A
43	277	17.7	4248	3	US-10-164-595-53	Sequence 53, App1
44	275	17.6	4056	3	US-10-164-595-55	Sequence 55, App1
45	272	17.4	6158	3	US-09-799-451-897	Sequence 897, App

ALIGNMENTS

RESULT 1
US-08-098-327E-32
Sequence 32, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: DRUJLHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-32

Alignment Scores:
Pred. No.: 1,89e-154 Length: 950
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-31 (1-316) x US-08-098-327E-32 (1-950)

QY 1 SerAspLeuGluGlnGluArgAlaValysGluLeuGlnGlnGlnSerAspLeu 20
DB 3 ACCGACTCTGAACCAAGAGACGCTGCTAAAGAAAGTTCCAGAACCAAGCGATTGA 62
QY 21 GluGlnAspArgLeuAlaValysGluLeuGlnGlnGlnSerAspLeuGlnGlu 40
DB 63 GAAACAAGATAGACTTCTTAAGAAAAGTTCAAGACACCAAGCATTAGAACAGAG 122
QY 41 ArgArgAlaValysGluLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgAla 60
DB 123 AGACTTGCTAAAGAAAGTTGCAAGAACCAAGCGCTGCTAAAGAGAGACGCTGCT 182
QY 61 ValysGluLeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaValys 80
DB 183 AAAAGAAAGTTGCAAGAACCAAGCGATTGAAACAGAGACGCTGCTAAAGAAAG 242
QY 81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnGlnAspArgLeuAlaValys 100
DB 243 TTGCAAGAACCAAGCGATTGAAACAGAGACGCTGCTAAAGAAAGTTGCAAGAG 302
QY 101 GlnGlnSerAspLeuGlnGlnGlnGlnGlnArgAlaValysGluLeuGlnGln 120
DB 303 CACCAAGCGATTGAAACAGAGACGCTGCTAAAGAAAGTTGCAAGAACCAAGAGC 362
QY 121 AspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
DB 363 GATTAGAACCAAGAGACGCTGCTAAAGAAAGTTGCAAGAACCAAGCGATTGAGAA 422
QY 141 GlnGlnArgArgAlaValysGluLeuGlnGlnGlnGlnGlnSerAspLeuGlnGln 160
DB 423 CAAAGAGACTTCTTAAGAAAAGTTGCAAGAACCAAGCGATTGAGAACAGAGAG 482
QY 161 ArgAlaValysGluLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 483 CGTCTAAAGAAAGTTGCAAGAACCAAGCGATTGAAACAGAGAGACGCTGCTAAAG 542
QY 181 GluLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 543 GAAAGATTGCAAGAACCAAGCGATTGAAACAGAGAGACGCTGCTAAAGAAAGTTG 602
QY 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 603 CAAAGAGACCAAGAGATTGAAACAGAGAGACGCTGCTGAAAAAAATTTAGAAAG 662
QY 221 ValysGluLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
DB 663 AAAAGGAAACATGAGATATATTAGCAGAGATTATGCTGCTTTAGAAATACCGCT 722

QY 241 IleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHleGlnSerSerLeuPro 260
DB 723 ATGAACTTCATCAGAAAATGAAAGCTGAGATTATATACACACCAATCTTCTTACT 782
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerLeysGluIleSerIleIleGluValThrAsn 280
DB 783 CAGGACAAACAGAGGAGATAGTAGAGATTCCAAAGAAATATCTATATAGAAAAACAAT 842
QY 281 ArgGluSerIleThrThrAsnValGluGluValArgAspIleHleValGlyHisLeuGlu 300
DB 843 AGAGATCTATTACAAAGATTTGAAGAGACCAAGGAGATATACATPAAAGACATCTTGA 902
QY 301 GluValysAspGlySerIleValProGluGlnValysGluAspLeysSer 316
DB 903 GAAAGAAAGATGCTTCAATTAACCAAGAACAAAGAAAGAAATCT 950

RESULT 2
US-08-462-625-32
Sequence 32, Application US/08462625
Parent No. 619502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Swecker & Machie
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22133-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 39,300
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-32

Alignment Scores:
Pred. No.: 1,89e-154 Length: 950
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0

Db 193 AAAAAAAAAAGTTCAGAACAAACAAAGCGATTTAGAACAGAGACGCTCTAAAGAAAAG 252
Qy 81 LeuGIngluGIngluSerleuGIngluGIngluGIngluGIngluGIngluGIngluGInglu 100
Db 253 TTGCAAGAAACAACAAAGCGATTTAGAACAGATGCTTAAAGAAAAGTTACAGAG 312
Qy 101 GIngluSerleuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 120
Db 313 CAGCAAAAGCGATTTAGAACAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAAC 372
Qy 121 AapleuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 140
Db 373 GATTTAGAACAGAGACGCTCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAA 432
Qy 141 GIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 160
Db 433 CAAGAGAGCTTCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAGAA 492
Qy 161 ArgAlaLyuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 180
Db 493 CCGTCTAAAGAAAAGTTGCAAGAACAAACGCGATTTGCAAGAACAGAGACGCTCTAA 552
Qy 181 GluLyuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 200
Db 553 GAAAAGTTCAGAACAAACAAAGCGATTTAGAACAGAGACGCTCTAAAGAAAAGTTG 612
Qy 201 GIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 220
Db 613 CAAGACGACCAAGATTTAGAACAAAGCGATTCGATTCGAAAAAATTTAGAAA 672
Qy 221 LysLyuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 240
Db 673 AAAAAAGCACTGAGATATATATAGCAAGAGATTTATAGCTCTTAAAGATCCAGCT 732
Qy 241 IleGluProSerGluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 260
Db 733 ATGAGACTTCATCAAGAAAATGAAACGTGATTTATATATATATATATATATATATATAT 792
Qy 261 GlnAapAapGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 280
Db 793 CAGGACCAAGAGGAGATGATGAGATTCAGAGAAAATATCTATATATATATATATATATAT 852
Qy 281 ArgGluSerIleThrAenValGluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 300
Db 853 AGGAGATCTATTCACAAATGCTTGAAGACAGAGATTTATATATATATATATATATATATAT 912
Qy 301 GluLyuLysAapGlySerIleLysProGluGIngluGIngluGIngluGIngluGInglu 316
Db 913 GAAAAAGAAAGATGCTCAATTAACAGAACAAAGAAAAGATTAATCT 960
RESULT 4
US-08-462-625-41
; Sequence 41, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPIPOPS CAPABLE
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/462, 625
;/ FILING DATE: 05-JUN-1995
;/ CLASSIFICATION: 424
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 08/098, 327
;/ FILING DATE: 24-NOV-1993
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: FR 91 01286
;/ FILING DATE: 05-PEB-1991
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: McGowan, Malcolm K.
;/ REGISTRATION NUMBER: 39,300
;/ REFERENCE/DOCKET NUMBER: 010830-078
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (703) 836-6620
;/ TELEFAX: (703) 836-2021
;/ INFORMATION FOR SEQ ID NO: 41:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1482 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: DNA (genomic)
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: 1..1482
;/ PUBLICATION INFORMATION:
;/ DOCUMENT NUMBER: WO 92/13884
;/ PUBLICATION DATE: 20-AUG-1992
;/ US-08-462-625-41
Alignment Scores:
Pred. No.: 3,366-154
Score: 1549.00
Percent Similarity: 99.37%
Best Local Similarity: 99.37%
Query Match: 99.10%
DB: 3 Gaps: 0
US-09-837-344-31 (1-316) x US-08-462-625-41 (1-1482)
Qy 1 SerAapLeuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 20
Db 13 AGCGATCTTAGAACAG 72
Qy 21 GluGlnAapArgLeuAlaLysGluLysLeuGIngluGIngluGIngluGIngluGInglu 40
Db 73 GAACAAAGATGAGCTTCTTAAAGAAAAGTTACAAAGACAGCAAGCGATTTAGAACAAAG 132
Qy 41 ArgArgAlaLysGluLysLeuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 60
Db 133 AACCTTGTAAAGAAAAGTTGCAAGAACAAACGCGATTCGAAACAAAGAGAGACGCT 192
Qy 61 LysGluLysLeuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 80
Db 193 AAGAGAAAGCTTCAAGAACAAACAAAGCGATTTAGAACAGAGAGAGAGAGAGAGAGAG 252
Qy 81 LeuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 100
Db 253 TTGCAAGAACAAACAAAGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTACAAAG 312
Qy 101 GIngluSerleuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 120
Db 313 CAGCAAAAGCGATTTAGAACAGAGACGCTTAAAGAAAAGTTGCAAGAACAAAC 372
Qy 121 AapleuGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 140
Db 373 GATTTAGAACAG 432
Qy 141 GIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGIngluGInglu 160
Db 433 CAAGAGAGCTTCTTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAGAGAG 492


```

Qy      261  Glnshpshnaagglvaensserfargapserfegluileserlleieglulvethrasn 280
Db      793  CAGGACACACAGAGGAGATGATGAGATTCCAAAGAAATGCTTATTAATGAAAAACAAAT 852
Qy      281  ArggluberierlerturthrasnvaIGluGlyVargArspIleHisLysGluLeuGlu 300
Db      853  AGAGATCTATTACACAAATCTTGAAGAGACGAAAGGAGTATACATAAGACATCTTGAA 912
Qy      301  GlulysylvaapGlySerIleLysProGluGlnLysGluapLysSer 316
Db      913  GAAAAAGAAAGATGTTCAATATAAACAGAACAAAGAAAGATTAATCT 960

RESULT 6
US-08-462-625-45
/ Sequence 45, Application US/08462625
/ Patent No. 6319502
/ GENERAL INFORMATION:
/ APPLICANT: GUERIN-MARCHAND, Claudine
/ APPLICANT: DRULHE, Pierre
/ TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
/ TITLE OF INVENTION: HEPATIC STAGES OF F. FALCIPARUM BEARING EPITOPES CAPABLE
/ TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
/ STREET: P.O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/462,625
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/098,327
/ FILING DATE: 24-NOV-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 91 01286
/ FILING DATE: 05-FEB-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McGowan, Malcolm K.
/ REGISTRATION NUMBER: 39,300
/ REFERENCE/DOCKET NUMBER: 010830-078
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 45:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1482 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..1482
/ PUBLICATION INFORMATION:
/ DOCUMENT NUMBER: WO 92/13884
/ PUBLICATION DATE: 20-AUG-1992
/ US-08-462-625-45

Alignment Scores:
Pred. No.: 6,97e-154 Length: 1482
Score: 1546.00 Matches: 313
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2

```

Query Match:	98.91%	Indel:	0
DB:	3	Gaps:	0
US-09-837-344-31 (1-316) x US-08-462-625-45 (1-1482)			
QY 1 SerAspLeuGInGInGInuAArgAArgAlaLyseGInuLyseGInGInGInGInserAspLeu 20			
DB 13 AGCGATCTAGAACAGAGACGCTCTAAAGAAAAGTTGCCAAGAACCAACAAAGCGATTTA 72			
QY 21 GInGInAspAArgLeuAlaLyseGInuLyseLeuGInGInGInGInserAspLeuGInGIn 40			
DB 73 GAACAAGATTAGCTTGTCTAAAGAAAAGTTTACAGAGACGAAAGCGATTAGAACAGAG 132			
QY 41 ArgAArgAlaLyseGInuLyseLeuGInGInGInGInserAspLeuGInGInGInuArgAAla 60			
DB 133 AGACTGCTTAAAGAAAAGTTGCCAAGAACCAACAAAGCGATTAGAACAGAGACGCTCT 192			
QY 61 LyseGInuLyseLeuGInGInGInGInserAspLeuGInGInGInuAArgAArgAlaLyseGInuLyse 80			
DB 193 AAAGAAAAGTTGCCAAGAACCAACAAAGCGATTAGAACAGAGACGCTCTAAAGAAAAG 252			
QY 81 LeuGInGInGInGInserAspLeuGInGInAspAArgLeuAlaLyseGInuLyseLeuGInGIn 100			
DB 253 TTGCAGAGAACCAAGAAAGCGATTAGAACAGATTAGCTTAAAGAAAAGTTACAGAG 312			
QY 101 GInGInserAspLeuGInGInGInuAArgAArgAlaLyseGInuLyseLeuGInGInGInGInser 120			
DB 313 CAGCAAGACGATTATAGAACAGAGACGCTCTAAAGAAAAGTTGCCAAGAACCAACAAAGC 372			
QY 121 AspLeuGInGInGInuAArgAArgAlaLyseGInuLyseLeuGInGInGInGInserAspLeuGIn 140			
DB 373 GATTATAGAACAGAGACGCTCTAAAGAAAAGTTGCCAAGAACCAACAAAGCGATTAGAA 432			
QY 141 GInGInuAArgAArgAlaLyseGInuLyseLeuGInGInGInGInserAspLeuGInGInuAArg 160			
DB 433 CAAGAGACGCTTGTCTAAAGAAAAGTTGCCAAGAACCAAGAAAGCGATTAGAACAGAGAG 492			
QY 161 ArgAlaLyseGInuLyseLeuGInGInGInGInserAspLeuGInGInGInuAArgAArgAlaLyse 180			
DB 493 CGTGTCTAAAGAAAAGTTGCCAAGAACCAACAAAGCGATTAGAACAGAGACGCTCTAA 552			
QY 181 GInuLyseLeuGInGInGInGInserAspLeuGInGInGInuAArgAArgAlaLyseGInuLyseLeu 200			
DB 553 GAAAAGTTGCCAAGAACCAACAAAGCGATTAGAACAGAGACGCTCTAAAGAAAAGTTG 612			
QY 201 GInGInGInGInuAArgAspLeuGInGInGInuArgLyseAArgThrLyseLyseAsnLeuGInuArg 220			
DB 613 CAAGAGACGCAAGAAAGATTAGAACCAAGAGAAAGCGATTAGCAAGAAAAGTTTAAAGAAAG 672			
QY 221 LyseLyseGInuLyseGInuAspLeuAlaLyseGInuArgLeuGInuLeuProAla 240			
DB 673 AAAAGAGAAACATGAGATATTATACAGAGAGATTTATATAGCTGCTTTAAAGAAAAGCAAGCT 732			
QY 241 ILeuGInuProSerGInuAsnGInuArgLyseGInuArgLyseProGInuLeuProGInGInGInser 260			
DB 733 ATAGAACTTCATTCAGAAAATAGACGTGATATATTATATACCATCAATCAATCTTTTAACT 792			
QY 261 GInAspAsnArgGInuAsnSerArgAspSerLyseGInuLeuSerLeuLeuGInuArgThrAsn 280			
DB 853 AAGAGATCTATTAACAACAATGTTGAGAGACCAAGAGATATCATTAAGACATCTTTGAA 852			
QY 301 GInuLyseLyseAspGInuSerLeuLyseProGInuLeuGInuAspLyseSer 316			
DB 913 GAAAAGAAAGATGTTCAATTAACCAAGAACAAAAGAAAGATTAATCT 960			

RESULT 7

US-08-098-327E-38

Sequence 38, Application US/08098327E

Patent No. 6270771

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-38

Alignment Scores:
Pred. No.: 2,47e-145 Length: 1493
Score: 1465.00 Matches: 311
Percent Similarity: 98.42% Conservative: 0
Best Local Similarity: 98.42% Mismatches: 5
Query Match: 93.73% Indels: 3
DB: Gaps: 0

US-09-837-344-31 (1-316) x US-08-462-625-38 (1-1493)

QY 1 SerApleuGluGlnGluArgGluAlaLysGluGluGlnGlnInserApleu 20
DB 13 AGCGATCTTGAAACAAAGAGAGAGCTGCTTAAGAAAGTTGCAAGAACAAAGCGATTGA 72
QY 21 GlnGlnAaPArgLeuAlaLysGluLysLeuGlnGlnInserApleuGlnGlu 40
DB 73 GAACAAGATAGACTTCTTAAGAAAGTTTACAAAGACACAAAGCGATTGACAAAGAG 132
QY 41 ArgAArgAlaLysGluLysLeuGlnGlnInserApleuGlnGlnGluArgAla 60
DB 133 AGACTTGTCTAA-GAAAAAGTTGCAAGAACAAAGCGATCTGAACAAAGAGAGAGCTGCT 191
QY 61 LysGluLysLeuGlnGlnInserApleuGlnGlnGluArgAlaLysGluLys 80
DB 192 AAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAAGAGAGAGCTGCTTAAGAAAG 251
QY 81 LeuGlnGlnGlnInserApleuGlnGlnAaPArgLeuAlaLysGluLysLeuGlnGlu 100
DB 252 TTGCAAGAACAAAGCGATTGAAACAAAGATAGACTTGTAAAGAAAGTTTACAAAGAG 311
QY 101 GlnGlnInserApleuGlnGlnGluArgAlaLysGluLysLeuGlnGlnGlnInser 120
DB 312 CAGCAAAAGCGATTGAAACAAAGAGAGAGCTGCTTAAGAAAGTTGCAAGAACAAAGAG 371
QY 121 AaPLeuGlnGlnGlnGluArgAlaLysGluLysLeuGlnGlnGlnInserApleuGln 140
DB 372 GATTTGAAACAAAGAGAGAGCTGCTTA-GAAAAAGTTGCAAGAACAAAGCGATTGAAAG 430
QY 141 GlnGlnAaPArgAlaLysGluLysLeuGlnGlnInserApleuGlnGlnGlnGluArg 160
DB 431 CAAGAGAGAGCTTGTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAAGAGAGAG 490
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnInserApleuGlnGlnGlnGluArgAlaLys 180
DB 491 CGTGCTTAAGAAAGAGTTGCAAGAACAAAGCGATTGAAACAAAGAGAGAGCTGCTTAAG 549
QY 181 GlnLysLeuGlnGlnGlnInserApleuGlnGlnGlnGluArgAlaLysGluLysLeu 200

DB 550 GAAAAAGTTGCAAGAACAAAGCGATTGAAACAAAGAGAGAGCTGCTTAAGAAAAAGTTG 609
QY 201 GlnGlnGlnGlnAaPArgLeuGlnGlnGluLysAlaAaPTrLysLysAaPLeuGlnGlu 220
DB 610 CAAGAGCAGCAAGAGAGATTGAAACAAAGAGAGCTGCTGATGCAAAAAAATTTAGAAAGA 669
QY 221 LysLysGlnGlnGlnAaPArgLeuAlaGluAaPLeuGluArgLysGluLysLeuAla 240
DB 670 AAAAGAGAACATGAGATATATTGACAGAGATTATATGTCGTTTGAATAACACACT 729
QY 241 IllegluePProSerGluAaPArgLysGluLysTrLysProHlAaPLeuSerLeuPro 260
DB 730 ATAGAACTTCATCAGAAAAAGAGCGTGAATATATATATACACATCAATCTTCTTACCT 789
QY 261 GlnAaPArgGluLysAaPArgAaPArgSerLysGluLysSerLysLysLysLysLys 280
DB 790 CAGGACAAACAGAGGAGATGAGATTCAGAGAAATATCTTATATGAAAAAACAAT 849
QY 281 ArgGlnSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 300
DB 850 AGAGATCTATTTACAAACAAAGTTGAAAGAGAGAGATATATCAATTAAGAGACATCTTGA 909
QY 301 GlnLysLysAaPArgLysLysLysLysLysLysLysLysLysLysLysLysLys 316
DB 910 GAAAAAGAAAGATGTTCAATTAACCAAGAACAAAAAGAAATTAATCT 957

RESULT 9
US-08-098-327E-37
Sequence 37, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: DRUIHE, Pierre
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..954


```
Db      1201 AGGAAAGAGCGAAGAAAGAAAGAGAGGAAAGAAAGAAAGAAAGAGAT 1260
Qy      44  LygGluLybLeuGInGInGInSerApeLeuGInGInuArGArGAlaLygGluLy 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1261 GAGAAAGAGAAAGAAAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAA 1320
Qy      64  LeuGInGInGInSerApeLeuGInGInuArGArGAlaLygGluLybLeuGInGlu 83
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1321 AGAGAGGAGCGGAAAGGAAA--GAGAGGAAAGGAGCGAGAAAGAAAGAAAGAA 1377
Qy      84  GInGInSerApeLeuGInGInApeArGLeuAlaLygGluLybLeuGInGInGInSer 103
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1378 GAGGAAAGAGAAAGAGAAAGAGAAAGAGCGAGAGGAAAGAAAGAGAGAAAGAA 1434
Db      104  ApeLeuGInGInGInuArGArGAlaLygGluLybLeuGInGInGInSerApeLeuGlu 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1435 -----GAGGAAAGAGCGAAAGAGAAAGAGAGAAAGAGAGAGAGAAAGAAAGAA 1488
Qy      124  GInGInuArGArGAla--LygGluLybLeuGInGInGInGInSerApeLeuGInGInGlu 142
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1489 GAGGAAAGAGAAAGATGAGGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAA 1548
Db      143  ArgArGAlaLybGluLybLeuGInGInGInSerApeLeuGInGInGInuArGArGAla 162
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1549 AGAAGAGAGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1608
Qy      163  LygGluLybLeuGInGInGInSerApeLeuGInGInuArGArGAlaLygGluLy 182
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1609 AAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1668
Db      183  LeuGInGInGInSerApeLeuGInGInuArGArGAlaLybGluLybLeuGInGlu 202
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1669 AGAGACGAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1728
Db      203  GInGInuArGArGLeuGInGIn-----ArgLybAlaApeThrLybLeuApeLeuGlu 219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1729 GAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1788
Qy      220  ArgLybLygGlu 223
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1789 AGAGAGAGAGAA 1800
```

Search completed: December 7, 2005, 12:45:24
Job time : 261.048 secs

THIS PAGE BLANK (USP19)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 10:33:33 ; Search time 1107.09 Seconds

(without alignments)
2360.343 Million cell updates/sec

Title: US-09-837-344-31

Perfect score: 1563
1 SDSEQERRAKEKQEQSD.....GHLEKKGSGIKPEQKEDKS 316

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Rgapop 6.0 , Rgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgn2_1/USPTO.spool/h/US09837344/runat.05122005.133535.839/app.query.fasta_1.718
-MAXLEN=2000000000 -USER=US09837344 @CGN 1.1 1364 @runat.05122005.133535.839
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1:	/cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
3:	/cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
4:	/cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
6:	/cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq:*
8:	/cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubna/US10F_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/1/pubna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	99.1	950	3	US-09-837-344-32
2	1549	99.1	950	3	US-09-900-963-32
3	1549	99.1	1482	3	US-09-837-344-41
4	1549	99.1	1482	3	US-09-900-963-41
5	1546	98.9	1482	3	US-09-837-344-45
6	1546	98.9	1482	3	US-09-900-963-45
7	1465	93.7	1493	3	US-09-837-344-38
8	1465	93.7	1493	3	US-09-900-963-38

9	771.5	49.4	954	3	US-09-837-344-37	Sequence 37, Appl
10	771.5	49.4	954	3	US-09-900-963-37	Sequence 37, Appl
11	771.5	49.4	988	3	US-09-837-344-34	Sequence 34, Appl
12	771.5	49.4	988	3	US-09-900-963-34	Sequence 34, Appl
13	369	23.6	3662	10	US-11-097-143-5152	Sequence 5152, Ap
14	329.5	21.1	4320	8	US-10-754-079-159	Sequence 159, App
15	327.5	21.0	3489	3	US-09-894-773-1	Sequence 1, Appl1
16	327.5	21.0	3489	6	US-10-294-804-1	Sequence 1, Appl1
17	327.5	21.0	3489	8	US-10-194-046-1	Sequence 1, Appl1
18	326	20.9	1835	3	US-09-216-393-80	Sequence 80, Appl
19	326	20.9	1835	6	US-10-321-856-80	Sequence 80, Appl
20	321	20.5	1662	10	US-11-097-143-5153	Sequence 5153, Ap
21	311	19.9	2479	6	US-10-108-260A-1966	Sequence 1966, Ap
22	307	19.6	720	5	US-10-101-487-74	Sequence 74, Appl
23	307	19.6	720	5	US-10-101-487-74	Sequence 74, Appl
24	307	19.6	720	9	US-10-939-988-76	Sequence 76, Appl
25	307	19.6	720	9	US-10-939-988-76	Sequence 76, Appl
26	306	19.6	7078	10	US-11-097-143-2762	Sequence 2762, Ap
27	306	19.6	13085	10	US-11-097-143-2761	Sequence 2761, Ap
28	294.5	18.8	1657	7	US-10-425-114-9663	Sequence 9663, Ap
29	294.5	18.8	2406	7	US-10-424-599-53216	Sequence 53216, A
30	292	18.7	42533	6	US-10-004-113-43	Sequence 43, Appl
31	291.5	18.7	1928	6	US-10-108-260A-1901	Sequence 1901, Ap
32	289	18.5	2772	9	US-10-450-763-8591	Sequence 8591, Ap
33	289	18.5	2772	9	US-10-450-763-26329	Sequence 26329, A
34	288	18.4	1824	9	US-10-450-763-17292	Sequence 17292, A
35	288	18.4	2093	6	US-10-104-047-1666	Sequence 1666, Ap
36	288	18.4	2850	9	US-10-450-763-15499	Sequence 15499, A
37	287	18.4	2243	7	US-10-276-774-433	Sequence 433, App
38	286.5	18.3	1522	6	US-10-104-047-1093	Sequence 1093, Ap
39	286.5	18.3	42999	3	US-09-799-462A-17	Sequence 17, Appl
40	286.5	18.3	42999	3	US-09-738-610A-17	Sequence 17, Appl
41	286.5	18.3	42999	3	US-09-738-610A-17	Sequence 73, Appl
42	286.5	18.3	42999	5	US-10-125-767-17	Sequence 17, Appl
43	286.5	18.3	42999	5	US-10-151-081-17	Sequence 17, Appl
44	286.5	18.3	42999	5	US-10-287-313-17	Sequence 17, Appl
45	286.5	18.3	42999	5	US-10-219-694-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-837-344-32
Sequence 32, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-Aug-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-837-344-32

Alignment Scores:
Pred. No.: 5,576-130 Length: 950
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-31 (1-316) x US-09-837-344-32 (1-950)

QY 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnInserAspLeu 20
DB 3 AGCGATCTGAACAAAGAGACCTGCTAAAGAAATTGCAAGAACAAAGCGATTGA 62
QY 21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlu 40
DB 63 GAACAAGATGAGACTTCTAAAGAAAGTTACAGACAGACAAAGCGATTGACAAAG 122
QY 41 ArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlnGlnGlnGln 60
DB 123 AGCTTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTGACAAAGAGACCTGCT 182
QY 61 LysGluLysLeuGlnGlnInserAspLeuGlnGlnGlnGlnGlnGlnGlnGln 80
DB 183 AAAGAAAGTTGCAAGAACAAAGCGATTGACAAAGAGACCTGCTAAAGAAAG 242
QY 81 LeuGlnGlnGlnInserAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGln 100
DB 243 TTGCAAGAACAAAGCGATTGACAAAGATGAGACTTCTAAAGAAAGTTACAAAG 302
QY 101 GlnGlnInserAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
DB 303 CAGCAAGCGATTGACAAAGAGACCTGCTAAAGAAAGTTGCAAGAACAAAGCG 362
QY 121 AspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
DB 363 GATTGAAACAGAGACCTGCTAAAGAAAGTTGCAAGAACAAAGCGATTGACAA 422
QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnInserAspLeuGlnGlnGln 160
DB 423 CAGGAGAGACTTCTAAAGAAAGTTGCAAGAACAAAGCGATTGACAAAGAGAG 482
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 483 CGGCTAAAGAAAGTTGCAAGAACAAAGCGATTGACAAAGAGACCTGCTAA 542
QY 181 GlnLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 543 GAAGAAGTTGCAAGAACAAAGCGATTGACAAAGAGACCTGCTAAAGAAAGTTG 602
QY 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 603 CAAAGAGACGAAAGAGATTGACAAAGAGAGCTGCTGACAAAGAAAGTTGACAA 662

QY 221 LysLysGlnHisGlyAspIleLeuAlaGluAspLeuTyrgIyArgLeuGluIleProAla 240
DB 663 AAAAGAAACATGAGATATATTAGCAGAGATTTATATGGTCGTTGAAATACAGCT 722
QY 241 IleGluLeuProSerGluAsnGluArgGlyTyrgIyIleProHisGlnSerSerLeuPro 260
DB 723 ATAGAACTTCATCAGAAATGAAACGTGATATATATACCATCAATCTCTTACCT 782
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThrAsn 280
DB 783 CAGGACAAACAGGGAGATGAGATTCAGAGAAATATCTATATATGAAAAACAAAT 842
QY 281 ArgGluSerIleThrThrAsnValGluIyArgArgAspIleHisLysGlyHisLeuGlu 300
DB 843 AGAGAATCTATTACACAAATCTTGAAGACGAAAGCATATACATTAAGACATCTTGA 902
QY 301 GlnLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 316
DB 903 GAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAAGATTAATCT 950

RESULT 2
US-09-900-963-32
Sequence 32, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUG/LIB: Piete
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIT C STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-Aug-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-900-963-32

Alignment Scores:
Pred. No.: 5,576-130 Length: 950

Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-31 (1-316) x US-09-900-963-32 (1-950)

```
OY 1 SerAspLeuGluGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeu 20
DB 3 AGGCATCTAGAACAAAGAGAGCGTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 62
OY 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlu 40
DB 63 GAACAAAGATAGACTTGTCTAAAGAAAGTTACAAAGAGCAAGCGATTAGAACAAAG 122
OY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGluArgAla 60
DB 123 AGACTTGCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGCGTGCT 182
OY 61 LysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGluArgArgAlaLysGlu 80
DB 183 AAAGAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGCGTGCTAAAGAAAG 242
OY 81 LeuGlnGlnGlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGln 100
DB 243 TTGCAAGAACAAAGCGATTAGAACAAAGATAGACTTGTCTAAAGAAAGTTACAAAG 302
OY 101 GlnGlnSerAspLeuGlnGlnGlnGluArgArgAlaLysGluLysLeuGlnGln 120
DB 303 CAGCAAGCGATTAGAACAAAGAGAGCGTGCTTAAGAAAGTTGCAAGAACAAAGCG 362
OY 121 AspLeuGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeu 140
DB 363 GATTAGAACAAAGAGAGCGTGCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAA 422
OY 141 GlnGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln 160
DB 423 CAAGAGAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAG 482
OY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgAla 180
DB 483 CGTGCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGCGTGCTTAA 542
OY 181 GluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnGluArgArgAlaLysGlu 200
DB 543 GAAAGTTGCAAGAACAAAGCGATTAGAACAAAGAGAGAGCGTGCTTAAGAAAGTTG 602
OY 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 603 CAAGAGCGAGAAAGATTAGAACAAAGAGAGCGTGCTTAAGAAAGAAATTTAGAA 662
OY 221 LysLysGluLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
DB 663 AAAAGAGAACATGGAGATATATAGCAGAGGATTTATGGTCGTTTGAAGATCCAGCT 722
OY 241 IleGlnLeuProSerGluAsnGluArgGlyTyrTyrIleProHISGlnSerSerLeu 260
DB 723 ATAGAACTTCATCGAAGAAATGAGCGTGATTTATATACCAATCATCTTTCTTACT 782
OY 261 GlnAspAsnArgLysAsnSerArgAspSerLysGlnIleIleGlnLysThrAsn 280
DB 783 CAGGCAACAGAGCGGCAATAGTAGAGATTCCAAAGGAAATATCTATATATGAAAAA 842
OY 281 ArgGlnSerLysThrThrAsnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 300
DB 843 AGAGAACTATATTAACAACAAAGTTGTAAGAGCAAGAGGATATACATTAAGGACATCTTGA 902
OY 301 GluLysLysAspGlnSerLysLysProGlnGlnGlnGlnGlnGlnGlnGlnGln 316
DB 903 GAAAGAAAGATGCTTCAATATAAACCAAGAACAAAGAAAGATTAATCT 950
```

RESULT 3

```
US-09-837-344-41
; Sequence 41, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRULHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1482 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1482
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: NO 92/13984
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-837-344-41
Alignment Scores:
Pred. NO.: 8.96e-130 Length: 1482
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
DB: 3 Gaps: 0
US-09-837-344-31 (1-316) x US-09-837-344-41 (1-1482)
OY 1 SerAspLeuGlnGlnGluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeu 20
DB 13 AGGCATCTAGAACAAAGAGAGCGTGCTAAAGAAAGTTGCAAGAACAAAGCGATTTA 72
OY 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln 40
DB 73 GAACAAAGATAGACTTGTCTAAAGAAAGTTACAAAGAGCAAGCGATTAGAACAAAG 132
OY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGluArgAla 60
```

DB 133 AGACTTGCTAAAGAAAAGTTGCAAGAACCAAAAGCATCTAGAACAAAGAGAGAGCTGCT 192
QY 61 LYSGLULYLSLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 80
DB 193 AAAAGAAAAGTTCCAAAGAACCAAAAGCGATTAGAACAAAGAGAGAGCTGCTAAAGAAAAG 252
QY 81 LEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 100
DB 253 TTGCAAGAACCAAAAGCGATTAGAACCAAAAGCATCTAGAACAAAGAGAGAGCTGCT 312
QY 101 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 120
DB 313 CAGCAAGAGCGATTGAAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGC 372
QY 121 AEPLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 140
DB 373 GATTTAGAACCAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTAGAA 432
QY 141 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 160
DB 433 CAAAGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTAGAACCAAGAGAGAG 492
QY 161 ARGALALYSGULYLSLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 180
DB 493 CGTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTAGAACCAAGAGAGAGCTGCTAA 552
QY 181 GIULYLSLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 200
DB 553 GAAAAGTTCCAAAGAACCAAAAGCGATTAGAACCAAGAGAGAGCTGCTAAAGAAAAGTTG 612
QY 201 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 220
DB 613 CAAAGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTAGAACCAAAAGTTG 672
QY 221 LYSGLULYLSLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 240
DB 673 AAAAGAAAAGTTGCAAGAACCAAAAGCGATTAGAACCAAGAGAGAGCTGCTAAAGT 732
QY 241 LLEGLULYLSLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 260
DB 733 ATAGAACCTTCCATCAAGAACCAAAAGCGATTAGAACCAAGAGAGAGCTGCTAAAGT 792
QY 261 GINAPASAPARGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 280
DB 793 CAGGACAAACAGAGAGAGATGATGAGATTCAGAGAAATATCTATATATAGAAAAACAAT 852
QY 281 ATSGGLUSERLLETHRAENVALIGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 300
DB 853 AGAGAACTTATTTACAAACAAATGTTGAAAGCAAGAGAGATTTACATTAAGAGACATCTTGA 912
QY 301 GIULYLSYAPSPGLYSERLLELYSPROGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 316
DB 913 GAAAAGAAAAGATGCTTCAATTAACCAAGAACCAAAAGAGATTAATCT 960
RESULT 4
US-09-900-963-41
Sequence 41, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-900-963-41
Alignment Scores:
Pred. No.: 8,966-130 Length: 1482
Score: 1549.00 Matches: 314
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.10% Indels: 0
Gaps: 0
DB: 3
US-09-837-344-31 (1-316) x US-09-900-963-41 (1-1482)
QY 1 SERAPLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 20
DB 13 AGCGATCTTAGAACCAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAGAGCGATT 72
QY 21 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 40
DB 73 GAAACAGATGACCTTGCTTAAGAAAAGTTACAAAGAGAGAGAGATTTAGAACAGAG 132
QY 41 ARGAGALALYSGULYLSLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 60
DB 133 AGACTTGCTAAAGAAAAGTTGCAAGAACCAAAAGCATCTAGAACCAAGAGAGAGCTGCT 192
QY 61 LYSGLULYLSLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 80
DB 193 AAAAGAAAAGTTCCAAAGAACCAAAAGCGATTAGAACCAAGAGAGAGCTGCTAAAGAAAAG 252
QY 81 LEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 100
DB 253 TTGCAAGAACCAAAAGCGATTAGAACCAAAAGCATCTAGAACCAAGAGAGAGCTGCT 312
QY 101 GINGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 120
DB 313 CAGCAAGAGCGATTGAAACAAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGC 372
QY 121 AEPLEUGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNG 140
DB 373 GATTTAGAACCAAGAGAGAGCTGCTAAAGAAAAGTTGCAAGAACCAAAAGCGATTAGAA 432

QY 141 GlnGluArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnArg 160
DB 433 CAAGAAGACTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGAGA 492
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnArgAlaLys 180
DB 493 CGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGAGCGTCTAA 552
QY 181 GluLysLeuGlnGlnGlnGlnSerAspLeuGlnGlnArgAlaLysGluLysLeu 200
DB 553 GAAAGAGTTGCAAGAACAAAGCGATTGAAACAAGAGCGTCTAAAGAAAGTTG 612
QY 201 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 613 CAAGAGCAGCAAAAGATTGAAACAAGAGCGTCTAAAGAAAGTTGAAAGA 672
QY 221 LysLysGluHISGlyAspIleLeuAlaGluAspLeuTyrGlyValGluGlnIleProAla 240
DB 673 AAAAGGAACTGGAGATATATTAGCAGAGGATTTATATGCTTAAAGATCCAGCT 732
QY 241 IleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHISGlnSerSerLeuPro 260
DB 733 ATAGACTTCATCAGAAATGAAAGTGTGATATTATATACCAATCAATCTTTTACT 792
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerIleSerIleIleGluLysThrAsn 280
DB 793 CAGACAAACAGAGGAATGATGAGATTCCAAAGAAATATCTATATATGAAAAACAAT 852
QY 281 ArgGluSerIleThrAsnValGluGlnValArgArgAspIleHISLysGluLysLeuGln 300
DB 853 AGAGATCTTATTAACAACAATGTTGAAAGCAAGAAAGGATATACATAAGGACATCTTGA 912
QY 301 GluLysLysAspGlySerIleLysProGlnGlnLysGluAspLysSer 316
DB 913 GAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAAGATTAATCT 960
RESULT 5
US-09-837-344-45
Sequence 45, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPIDIC SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Machis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22213-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-837-344-45
Alignment Scores:
Pred. No.: 1,67e-129 Length: 1482
Score: 1546.00 Matches: 313
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 98.91% Indels: 0
Gaps: 0
US-09-837-344-31 (1-316) x US-09-837-344-45 (1-1482)
QY 1 SerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 20
DB 13 AGCGATTAGAACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGA 72
QY 21 GluGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGln 40
DB 73 GAACAAATATGACTTGTCTAAAGAAAGTTTCAAGAGCAGCAAGCGATTGAAACAAG 132
QY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
DB 133 AGACTTCTTAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGAGCGTCT 192
QY 61 LysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
DB 193 AAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGAGCGTCTAAAGAAAG 252
QY 81 LeuGln 100
DB 253 TTGCAAGAACAAAGCGATTGAAACAAGATGACTTCTAAAGAAAGTTTCAAGAG 312
QY 101 GlnGlnSerAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
DB 313 CAGCAAGCGATTTTGAACAAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCG 372
QY 121 AspLeuGln 140
DB 373 GATTTAGAACAAAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGAA 432
QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
DB 433 CAAGAGCAGCTTGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGAGAGA 492
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 493 CGTCTAAAGAAAGTTGCAAGAACAAAGCGATTGAAACAAGAGCGTCTAA 552
QY 181 GluLysLeuGln 200
DB 553 GAAAGAGTTGCAAGAACAAAGCGATTGAAACAAGAGCGTCTAAAGAAAGTTG 612
QY 201 Gln 220
DB 613 CAAGAGCAGCAAAAGATTGAAACAAGAGCGTCTAAAGAAAGTTTGAAGA 672
QY 221 LysLysGluHISGlyAspIleLeuAlaGluAspLeuTyrGlyValGluGlnIleProAla 240

Db 673 AAAAAGAAACATGAGATATATTACAGAGATTTATATGCGTTGGAATAACACT 732
Qy 241 TleGIuLeuProSeRGlubnGluArglyTyTyTlleProHleGISeRSeuPro 260
Db 723 ATAGAACTTCATCAGAAAATGACGTGATATATATACCAATCAATCTTCTTACT 792
Qy 261 GlnAaPaenArgGlyAaenSerArgAaSerLySeRlySeRlleGIleGIuLeuThraen 280
Db 793 CAGGACACAGAGGAAATGAGATTCAGAGAAATGCTATATATGAAAAACAAT 852
Qy 281 ArgGlySeRlleThraenValGluGlyArgArgAaPleHleIleLySeRlyHleuGlu 300
Db 853 AGAGATCTATTACAAACAATGTTGAGAGCAGAGGATATCATTAAGACATCTTGA 912
Qy 301 GlnLybLyAaPlySeRlleLyProGluGlnLybGluAaPlySeR 316
Db 913 GAAAAGAAAGATGTTCAATAAACCAAGAAAAAGAAATAAATCT 960

RESULT 6
US-09-900-963-45
Sequence 45, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-900-963-45

Alignment Scores:

	Pred. No.:	1.67e-129	Length:	1482
	Score:	1546.00	Matches:	313
	Percent Similarity:	99.37%	Conservative:	1
	Best Local Similarity:	99.05%	Mismatches:	2
	Query Match:	98.91%	Indels:	0
	DB:	3	Gaps:	0
US-09-837-344-31 (1-316) x US-09-900-963-45 (1-1482)				
Qy	1	SeRleuGluGlnGluArgAlaLySeGluLySeuGlnGlnGlnSeRleu	20	
Db	13	AGCGATTCAGAACAAAGAGACGTCTAAAGAAAAGTTGCAGAACAAAGCATTTA	72	
Qy	21	GlnGlnAaPlyLeuAlaLySeGluLySeuGlnGlnGlnSeRleuGlnGlu	40	
Db	73	GACAAAGTACCTTGCTTAAAGAAAAGTTACAGAGCAGAAAGCATTTGAAACAAG	132	
Qy	41	ArgArgAlaLySeGluLySeuGlnGlnGlnSeRleuGlnGlnArgAla	60	
Db	133	AGACTTGCTTAAAGAAAAGTTGCAGAACAAAGCATTCAGAACAAAGACGTGCT	192	
Qy	61	LySeGluLySeuGlnGlnGlnSeRleuGlnGlnArgArgAlaLySeGluLyS	80	
Db	193	AAAGAAAAGTTGCAGAACAAAGCATTCAGAACAAAGACGTGCTTAAAGAAAAG	252	
Qy	81	LeuGlnGlnGlnSeRleuGlnGlnAaPlyLeuAlaLySeGluLySeuGlnGlu	100	
Db	253	TTGCAGAACAAAGCATTCAGAACAAAGCATTCGTTAAAGAAAAGTTACAGAG	312	
Qy	101	GlnGlnSeRleuGlnGlnGlnArgArgAlaLySeGluLySeuGlnGlnGlnSeR	120	
Db	313	CAGCAAAAGCATTCAGAACAAAGCATTCGTTAAAGAAAAGTTGCAGAACAAAGC	372	
Qy	121	AaPlyLeuGlnGlnGlnArgArgAlaLySeGluLySeuGlnGlnGlnSeRleuGlu	140	
Db	373	GATTTAGAACAGAGACGTCTAAAGAAAAGTTGCAGAACAAAGCATTTAGAA	432	
Qy	141	GlnGlnArgArgAlaLySeGluLySeuGlnGlnGlnSeRleuGlnGlnGlnArg	160	
Db	433	CAAGAGACATTCGTTAAAGAAAAGTTGCAGAACAAAGCATTTAGAACAGAGAG	492	
Qy	161	ArgAlaLySeGluLySeuGlnGlnGlnSeRleuGlnGlnGlnArgArgAlaLyS	180	
Db	493	CGTGCTTAAAGAAAAGTTGCAGAACAAAGCATTCGTTAGAACAGAGAGAGTGCTTAA	552	
Qy	181	GluLySeuGlnGlnGlnSeRleuGlnGlnGlnArgArgAlaLySeGluLySeu	200	
Db	553	GAAAAGTTGCAGAACAAAGCATTCAGAACAAAGACGTGCTTAAAGAAAAGTTG	612	
Qy	201	GlnGlnGlnArgArgleuGlnGlnArgLyAlaSeRThraLySeRleuGlnArg	220	
Db	613	CAAGACACAGAGATTTAGAACAAAGGATGATGCAAAAAAATTTAGAAAGA	672	
Qy	221	LybLySeGluHleGIleuAlaLySePlyTyTyTyArgleuGluTyTyTyProAla	240	
Db	673	AAAAAGAAACATGAGATATATTACAGAGGATTTATATGTCGTTGAAAAATCACAGCT	732	
Qy	241	lleGIuLeuProSeRGlubnGluArglyTyTyTlleProHleGISeRSeuPro	260	
Db	733	ATAGAACTTCATCAGAAAATGACGTGATATATATACCAATCAATCTTCTTACT	792	
Qy	261	GlnAaPaenArgGlyAaenSerArgAaSerLySeRlySeRlleGIleGIuLeuThraen	280	
Db	793	CAGGACACAGAGGAAATGAGATTCAGAGAAATGCTATATATGAAAAACAAT	852	
Qy	281	ArgGlySeRlleThraenValGluGlyArgArgAaPleHleIleLySeRlyHleuGlu	300	
Db	853	AGAGATCTATTACAAACAATGTTGAGAGCAGAGGATATCATTAAGACATCTTGA	912	
Qy	301	GlnLybLyAaPlySeRlleLyProGluGlnLybGluAaPlySeR 316		
Db	913	GAAAAGAAAGATGTTCAATAAACCAAGAAAAAGAAATAAATCT 960		


```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/900,963
  FILING DATE: 10-Jul-2001
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/098,327
  FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
  NAME: McGowan, Malcolm K.
  REGISTRATION NUMBER: 39,300
  REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (703) 836-6620
  TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1493 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
  DOCUMENT NUMBER: WO 92/13884
  PUBLICATION DATE: 20-Aug-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-900-963-38

Alignment Scores:
  Pred. No.: 3,2e-122      Length: 1493
  Score: 1465.00          Matches: 311
  Percent Similarity: 98.42%      Conservative: 0
  Best Local Similarity: 98.42%      Mismatches: 5
  Query Match: 93.73%          Indels: 3
  DB: 3                      Gaps: 0

US-09-837-344-31 (1-316) x US-09-900-963-38 (1-1493)
QY 1 SerAspLeuGluGlnGluArgAlaLysGluLysLeuGlnGlnGlnSerAspLeu 20
DB 13 ACGGATCTAGAACAAAGAGACGCTCTTAAAGAAAGTTGCAAGAACAAAGCCGATTTA 72
QY 21 GlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlu 40
DB 73 GAACCAAGATGAGCTTCTAAGAAAGTTTACAAAGACGCAAGCGATTGACAAAG 132
QY 41 ArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgAla 60
DB 133 AGACTTGCTTAA-GAAAGTTGCAAGAACAAAGCGATCTTAAACAAAGAGACCTGCT 191
QY 61 LysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLysGluLys 80
DB 192 AAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAAAGAGACGCTCTAAAGAAAG 251
QY 81 LeuGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLysGluLysLeuGlnGlu 100
DB 252 TTGCAAGAACAAAGCCGATTGAAACAAAGATGACCTTAAAGAAAGTTTACAAAG 311
QY 101 GlnGlnSerAspLeuGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnSer 120
DB 312 CAGCAAGCGATTTAAGAACAGAGACGCTCTTAAAGAAAGTTGCAAGAACAAAGC 371
QY 121 AspLeuGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlu 140
DB 372 GATTTAGAACAAAGAGACGCTCTAA-GAAAGTTGCAAGAACAAAGCCGATTGAGA 430
QY 141 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGln 160
DB 431 CAAGAGAGACTTCTCTTAAAGAAAGTTGCAAGAACAAAGCGATTTGACAAAGAG 490
QY 161 ArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLys 180

```

```

DB 491 CGTGCTAAGAAAGTTGCAAGAACAAAGCCGATTGAAACAAAGAGACGCTCTAA- 549
QY 181 GlnLysLeuGlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLysGluLysLeu 200
DB 550 GAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCTCTTAAAGAAAGTTG 609
QY 201 GlnGlnGlnGlnArgAspLeuGlnGlnGlnArgAlaLysGluLysLeuGlnGlu 220
DB 610 CAGAGCAGCAAGAGATTTTGAACAAAGGAGCTGATACGAAAAAAATTGAAAGA 669
QY 221 LysLysGlnHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGlnIleProAla 240
DB 670 AAAAGGAACATGAGATATATTAGCAGAGGATTTATATGCTGTTGAAATACAGCT 729
QY 241 IleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerSerLeuPro 260
DB 730 ATAGAACTTCATCGAAGAAATGAAACGTGATTTATATACCATCAATCTTCTTACCT 789
QY 261 GlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThrAsn 280
DB 790 CAGGACCAACAGAGGAAATGATGAGATTTCCAAAGAAATATCTATATGAAAAACAAT 849
QY 281 ArgGlnSerIleThrThrAsnValGluGlyArgArgAspIleHisLysGlyHisLeuGlu 300
DB 850 AGAGAACTTATTAACAAGATGTTGAAAGACGAAAGGATATACATTAAGAGACATCTTGA 909
QY 301 GlnLysLysAspGlySerIleLysProGlnGlnLysGluAspLysSer 316
DB 910 GAAAGAAAGATGTTCAATTAACCAAGAACAAAGAAAGATTAATCT 957

RESULT 9
US-09-837-344-37
Sequence 37, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSER: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs

```

```

:
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 1..954
: PUBLICATION INFORMATION:
:   DOCUMENT NUMBER: WO 92/13884
:   PUBLICATION DATE: 20-AUG-1992
: SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-837-344-37

Alignment Scores:
Pred. No.: 4.17e-60 Length: 954
Score: 771.50 Matches: 163
Percent Similarity: 92.59% Conservative: 12
Best Local Similarity: 86.24% Mismatches: 13
Query Match: 49.36% Indels: 1
DB: Gaps: 1

US-09-837-344-31 (1-316) x US-09-837-344-37 (1-954)
QY 2 Aspleuglugingluarqargalalyegluylvseugluginglinseraepleuglu 21
DB 391 GATACGATTAATAAATAATATTATTAAGCGCAAGCAACAAACAGCAAGATCTTGAA 450
QY 22 Glnaapargleualalyegluylvseugluginglinseraepleuglugingluarq 41
DB 451 ---GAAAAGCAGCTAAAGAAAAGTTACAGGGGCAACAAACGATTCAGAACAGAGAGA 507
QY 42 Argalalyegluylvseugluginglinseraepleuglugingluarqargalalye 61
DB 508 CGTCTAAAGAAAAGTTGCAAGAACAAACGATTTAGAACAAAGAGAGACTTGCTAA 567
QY 62 Glulvseugluginglinseraepleuglugingluarqargalalyegluylvseu 81
DB 568 GAAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGACTTGCTAAAGAAAAGTTG 627
QY 82 Glngluginglinseraepleugluginglnaapargleualalyegluylvseuglugin 101
DB 628 CAAGACGATTTAGAACAGAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGAT 687
QY 102 Glnseraepleuglugingluarqargalalyegluylvseugluginglinserae 121
DB 688 CAAGCGATTTAGAACAGAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGAT 747
QY 122 Leuglugingluarqargalalyegluylvseugluginglinseraepleuglugin 141
DB 748 TTAGAACAAAGAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAA 807
QY 142 Gluargargalalyegluylvseugluginglinseraepleuglugingluarqarg 161
DB 808 GAGAGACTTGCTAAAGAAAAGTTACAGAGAGAGAGCGATTTAGAACAGATAGACTT 867
QY 162 Alalyegluylvseugluginglinseraepleuglugingluarqargalalyeglu 181
DB 868 GCTAAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGAGACTTGCTAAAGAAA 927
QY 182 Lylvseugluginglinseraepleu 190
DB 928 AGTTGCAAGAACAAACGATTTA 954

RESULT 10
US-09-900-963-37
: Sequence 37, Application US/09900963
: Publication No. US20030064075A1
: GENERAL INFORMATION:
:   APPLICANT: GUERIN-MARCHAND, Claudine
:   DRUILHE, Pierre
:   TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
:   HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
:   OF STIMULATING THE T LYMPHOCYTES
```

```

:
: NUMBER OF SEQUENCES: 46
: CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Burne, Doane, Swecker & Mathis
:   STREET: P.O. Box 1404
:   CITY: Alexandria
:   STATE: Virginia
:   COUNTRY: United States
:   ZIP: 22313-1404
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/900,963
:   FILING DATE: 10-Jul-2001
:   CLASSIFICATION: <Unknown>
:   PRIORITY APPLICATION DATA:
:     APPLICATION NUMBER: 08/098,327
:     FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
:   NAME: McGowan, Malcolm K.
:   REGISTRATION NUMBER: 39,300
:   REFERENCE/DOCKET NUMBER: 010830-045
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: (703) 836-6620
:   TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID NO: 37:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 954 base pairs
:     TYPE: nucleic acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
:     MOLECULE TYPE: DNA (genomic)
:     FEATURE:
:       NAME/KEY: CDS
:       LOCATION: 1..954
:     PUBLICATION INFORMATION:
:       DOCUMENT NUMBER: WO 92/13884
:       PUBLICATION DATE: 20-AUG-1992
:     SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-837-344-31 (1-316) x US-09-900-963-37 (1-954)
QY 2 Aspleuglugingluarqargalalyegluylvseugluginglinseraepleuglu 21
DB 391 GATACGATTAATAAATAATATTATTAAGCGCAAGCAACAAACAGCAAGATCTTGAA 450
QY 22 Glnaapargleualalyegluylvseugluginglinseraepleuglugingluarq 41
DB 451 ---GAAAAGCAGCTAAAGAAAAGTTACAGGGGCAACAAACGATTCAGAACAGAGAGA 507
QY 42 Argalalyegluylvseugluginglinseraepleuglugingluarqargalalye 61
DB 508 CGTCTAAAGAAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGAGACTTGCTAA 567
QY 62 Glulvseugluginglinseraepleuglugingluarqargalalyegluylvseu 81
DB 568 GAAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGAGACTTGCTAAAGAAAAGTTG 627
QY 82 Glngluginglinseraepleugluginglnaapargleualalyegluylvseuglugin 101
DB 628 CAAGACGATTTAGAACAGAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAAGCGAT 687

Alignment Scores:
Pred. No.: 4.17e-60 Length: 954
Score: 771.50 Matches: 163
Percent Similarity: 92.59% Conservative: 12
Best Local Similarity: 86.24% Mismatches: 13
Query Match: 49.36% Indels: 1
DB: Gaps: 1
```

```

OY      102  GINSERAEPLLEUGLNGLNGLUARGAALALYSGULYLEUGLNGLNGLNGINSERAP 121
DB      668  CAAACGGATTGAAACAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAACAGCGAT 747
OY      122  LEUGLNGLNGLUARGAALALYSGULYLEUGLNGLNGLNGINSERAEPLLEUGLNG 141
DB      748  TTAGAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAAACAGCGATTAGAACAA 807
OY      142  GLUATGATGALALYSGULYLEUGLNGLNGLNGINSERAEPLLEUGLNGLUATGATG 161
DB      808  GAGAGACTTGCTAAAGAAAAGTTACAAAGACGACAAAGCGATTAGAACAAAGTACCTT 867
OY      162  ALALYSGULYLEUGLNGLNGLNGINSERAEPLLEUGLNGLUARGAALALYSGUL 181
DB      868  GCTAAAGAAAAGTTGCAAGAACAAACAGCGATTGAAACAGAGAGACGCTGTTAAGAA 927
OY      182  LYLLEUGLNGLNGLNGINSERAEPLLEUGLNGLUARGAALALYSGUL 190
DB      928  AGCTTCCAAAGAACAAACAGCGATTTA 954

RESULT 11
US-09-837-344-34
; Sequence 34, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUIHE, Pierre
; TITLE OF INVENTION:
;   PEPTIDE SEQUENCES SPECIFIC FOR THE
;   HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
;   OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 988 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-837-344-34
Alignment Scores:

```

```

Pred. No.:      4,32e-60          Length:      988
Score:          771.50           Matches:       163
Percent Similarity: 92.59%       Conservative: 12
Best Local Similarity: 86.24%    Mismatches:   13
Query Match:     49.36%         Indels:        1
DB:              3               Gaps:          1

US-09-837-344-31 (1-316) x US-09-837-344-34 (1-988)

QY      2 Aspleuglucingluarargalalyegluylvleuglucinglucinsereaspleuglu 21
Dd      423 GATGACGATATAAAAAAATATATTAAAGGGCAAGACGAACAACAGAGAGTCTTGAA 482
QY      22 Glmasparglucumalyegluylvleuglucinglucinsereaspleuglucingluar 41
Dd      483 ---GAAAACACGTAAAGAAAAGTTACAGGGGCACAAAGCCATTTCAGAACAGAGAA 539
QY      42 Argalalyegluylvleuglucinglucinsereaspleuglucingluarargalalyeg 61
Dd      540 CGTCCTTAAGAAAAGTTGCCAAGAACAACAAGCGATTGTGAACAAGAGACTTGCTTAA 599
QY      62 Glulvleuglucinglucinsereaspleuglucingluarargalalyegluylvleu 81
Dd      600 GAAAAGTTGCAGAACAAACAAGCGATTTAGAACAAAGACAGCCTCTAAAGAAAAGTTG 659
QY      82 Glnglucinglucinsereaspleuglucinaeparglucumalyegluylvleuglucinglu 101
Dd      660 CAAACAACAACAAGCCGATTTAGAACACAAGACCTGCTTAACAAANAGTTGCAGAACAA 719
QY      102 Glinseraspleuglucingluarargalalyegluylvleuglucinglucinsereasp 121
Dd      720 CAAAGCGATTTAGAACACAAGACAGCCTGCTTAAGAAAAGTTGCAGAACACAACAAGCAT 779
QY      122 leuglucingluarargalalyegluylvleuglucinglucinsereaspleuglucin 141
Dd      780 TTAGAACAAAGAGAGACCTGCTTAAGAAAAGTTCCAAGAACACAACAAGCGATTTAGAACAA 839
QY      142 Gluarargalalyegluylvleuglucinglucinsereaspleuglucingluararg 161
Dd      840 GAGGAGCTTGCTTAAGAAAAGTTACAAAGACAGCAAAGCGATTGTAGAACAAAGTAGACTT 899
QY      162 Alalyegluylvleuglucinglucinsereaspleuglucingluarargalalyeglu 181
Dd      900 GCTTAAGAAAAGTTGCCAAGAACAACAAGCGATTTGAACAAGAGACCTGCTTAAAGAA 959
QY      182 Lysleuglucinglucinsereaspleu 190
Dd      960 AGTTTGCAGAACAAACAAGCGATTTA 986

RESULT 12
US-09-900-963-34
; Sequence 34, Application US/09900963
; Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATITIS STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22133-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```



```

Oy      124 GlnGluAaGaAGla---LybGlUuYbLeuGInGInGInGInSerAerLeuGInGInu 142
           ::::::::::::::::::::
Db      1174 GAGAAAGAAGGAAGATGACGAGAAAGAAAGAGAAAGGAAAAAGAAAAGAAAGAAAGAA 1115
Oy      143 ArgArGAlaLysGLuLYbLeuGInGInGInGInSerAerLeuGInGInuArGrAa 162
           1114 AGAAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1055
Db      163 LybGLuLYbLeuGInGInGInGInSerAerLeuGInGInuArGrAALaLYbGLuYs 182
           1054 AAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 995
Oy      183 LeuGInGInGInSerAerLeuGInGInuArGrAALaLYbGLuLYbLeuGInGInu 202
           994 AGAGACGAGAAACGAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAA 935
Oy      203 GInGIn-----ArgAerLeuGInGInuArGrYLaAerThrLYbLYbAerLeuGInuArG 220
           934 GAGCAAAAAAGAAAGCAAGCAAGCAAGCAAGAAAGAAAGAAAGCAAGCAAGCAAGCAAGCA 875
Oy      221 LYbLYbGLuHbGLYAerLleLeuAlaGLuAerLeuTYrGLYAarLeuGInuLleProAla 240
           Db      874 AAAAGAAAAAGAAAGCAGCAGCAAGAAAAAGACAAGCAAGAAAGCAAGCAAGCAAG-- 824
Oy      241 lLeglueProsErGluAbnGInuArGrGLuTYrTYrLleProH:lEGInSerSetleuPro 260
           Db      823 ---GAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 773
Oy      261 GlNaPeAmArGrGLYAserSerArGrAerSerlySGluIleSerllelEGluLYbArGrAn 280
           Db      772 GAGCAAAAAAGAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 722
Oy      281 ArGrGLuSerlleThrThraenValGLuGLYAaGrAerlleHieLYbGLYhiSleuGlu 300
           Db      721 AAAGCGAGAGAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 662
Oy      301 GlULyLYbAerPGlySerlleYsPrOGInuLYbGLUAerPYs 315
           Db      661 GAAAAAGCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 617

RESULT 14
US-10-754-079-159
/ Sequence 159, Application US/10754079
/ Publication No. US2005008602A1
/ GENERAL INFORMATION:
/ APPLICANT: Hiral, Yohe;
/ APPLICANT: Takebe, Kyoko
/ TITLE OF INVENTION: OLIGOPEPTIDES FOR PROMOTING HAIR GROWTH
/ FILE REFERENCE: 4673920022
/ CURRENT APPLICATION NUMBER: US/10/754, 079
/ CURRENT FILING DATE: 2004-01-07
/ PRIOR APPLICATION NUMBER: US 10/155, 922
/ PRIOR FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: PCT/JPO1/04691
/ PRIOR FILING DATE: 2001-06-04
/ PRIOR APPLICATION NUMBER: JP 2000-166903
/ PRIOR FILING DATE: 2000-06-05
/ PRIOR APPLICATION NUMBER: JP 2001-347340
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: JP 2001-347338
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: JP 2001-371175
/ PRIOR FILING DATE: 2001-12-05
/ PRIOR APPLICATION NUMBER: JP 2001-371336
/ PRIOR FILING DATE: 2001-12-05
/ PRIOR APPLICATION NUMBER: JP 2003-1891
/ PRIOR FILING DATE: 2003-01-08
/ NUMBER OF SEQ ID NOS: 162
/ SOFTWARE: FaastSeq for windows Version 4.0
/ SEQ ID NO 159
/ LENGTH: 4320
/ TYPE: DNA

```

[illegible]

```
OY 261 GluAaPnaAArgIYAaSerArGAaPSeRlySeGluIleSeRIleIleGluLyThAaSn 280
Db 2674 GAAGAGCAAGCGTGGCAAGAGCGGACCGTTTGAAGGAGGAGCGACCTTCCGCGACAG 2733
OY 281 ArgGlu-----SerIleThrThAaPnaIglu 289
Db 2734 CAGAGAGAAAGACAGCCCGCCCGCAGGAATGTGACAGAAATTAACCGGAACTAGAG 2793
OY 290 GLYAArgAArgApIleHleYSeGluYleuGluGluYleYSeRlySeRIleYPro 309
Db 2794 GTCCGCGCAGAACTGAGAAAGACGCGCTCGCGACAGAAAG-----CTCCGCGAG 2844
OY 310 GluGluYleY 312
Db 2845 GAACAAAG 2853

RESULT 15
US-09-894-273-1
; Sequence 1, Application US/09894273
; Publication No. US20040037847A1
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballester, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/894,273
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/109,422
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-1

Alignment Scores:
Pred. No.: 1,33e-19 Length: 3489
Score: 327.50 Matches: 79
Percent Similarity: 58.92% Conservative: 106
Best Local Similarity: 25.16% Mismatches: 104
Query Match: 20.95% Indels: 25
DB: 3 Gaps: 9

US-09-837-344-31 (1-316) x US-09-894-273-1 (1-3489)
OY 4 GluGluGluAArgAlaYSeGluYleuGluGluGluGluGluGluGluGluGluAaP 23
Db 1915 GAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAGAT 1974
OY 24 ArgLeuAlaYSeGluYleuGluGluGluGluGluGluGluGluGluGluGluGluGlu 39
Db 1975 GAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAGAT 2034
OY 40 GluAArgAlaYSeGluYleuGluGluGluGluGluGluGluGluGluGluGluGluGlu 59
Db 2035 GAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAGAT 2094
OY 60 AlaYSeGluYleuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 79
Db 2095 GAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAG 2154
OY 80 LySeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 99
Db 2155 CAGCAGATGAGCAGACAGCAGACAGCAGATGAGCAGACAGCAGATGAGCAGACAG 2214
OY 100 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 119
Db 2215 GATGAGCAGCAGCAGCAGATGAGCAGACAGCAGCAGATGAA-----CAGCAGACAG 2268
```

```
OY 120 SerASeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 139
Db 2269 GAGAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2328
OY 140 GluGluGluAArgAlaYSeGluYleuGluGluGluGluGluGluGluGluGluGluGlu 159
Db 2329 GAGGATCAG-----GAGCAGAGATTAGAGAGCAGCAGCAGCAGCAGCAGCAGCAG 2379
OY 160 ArgAArgAlaYSeGluYleuGluGluGluGluGluGluGluGluGluGluGluGluGlu 179
Db 2380 -----GAGCAGAGATTAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2421
OY 180 LySeGluYleuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 199
Db 2422 GAGCAGAGATTAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2472
OY 200 LeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 219
Db 2473 TTAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2523
OY 220 ArgLySeGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 239
Db 2524 CAGAGCGTGAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2583
OY 240 AlaIleGluLeuProSerGluAaSerGluAArgGlyTyTrIleProHISGluSerSeu 259
Db 2584 GAATTAGAGAGGTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2643
OY 260 ProGluAaPnaAArgIYAaSerArGAaPSeRlySeGluIleSeRIleIleGluYThr 279
Db 2644 GAGAGCGTGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2703
OY 280 AaAArgGluSerIleThrThAaPnaIgluGluYArgAArgApIleHleYSeGluYleu 299
Db 2704 GAGTTAGAG-----GAGGTGAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2751
OY 300 GluGluYleYSeGluYSeRlySeRIleYProGluGluGluGlu 313
Db 2752 GAAGAGCAGAGCAGCAGCGGCTGAAACAGCAGCAGCAGAG 2793
```

Search completed: December 7, 2005, 16:15:13
Job time : 1122.09 secs

THIS PAGE BLANK (USPTO)

QY 134 GlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGln 153
Db 1960 TTGTTTAATTCAGTAAGTAATGATGCGTTAAAGATGCTTTTGAAGGACCA--- 2016
QY 154 SerAspLeuGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGlnSerAspLeu 173
Db 2017 -----ATTGACGACACAGAAATGTAGACAAATTTAAATCAACAATTAGTGAAGATT 2070
QY 174 GlnGlnGlnArgArg---AlaLysGluLysLeuGlnGlnGlnSerAspLeuGlu--- 191
Db 2071 AAACAGCAGCTTAAGATCTTAATGAGAAATCGAAATTAATGAAGTCAGCTTCAAAA 2130
QY 192 -----GlnGlnArgArgAlaLysGlu--- 198
Db 2131 TGTGATCAGATATACCTTTCTATAGAAAATCATTTATCAAGATTTAAAGCAAACTCA 2190
QY 199 -----LysLeuGln 201
Db 2191 AAGCTAGATGATTAAATCAACCATGCAATAGATCATTTAAATGACAGCTATCAACTCA 2250
QY 202 GlnGlnGlnArgAspLeuGlnGlnArgLysAlaAspThrLysLysAsnLeuGlnArgLys 221
Db 2251 GTAGACCTGCAAGATGCAATGATCTGATGAAACTATTGCAATTTGCCGTAACAAA 2310
QY 222 LysGluLysGluAspLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 241
Db 2311 GTAAATTTAAACGAATGATGCAATGATGATTTAA---GGTCTGTAATTTAAAGCAATA 2367
QY 242 GluLeuProSerGlu---AsnGlnArgGlyTyrTyrIleProHisGlnSerSerLeuPro 260
Db 2368 GAACAATTGGAAGGTTGAATGAACGATACACATTTTAAATGAGCAA----- 2415
QY 261 GlnAspAsnArgLysAsnSerArgAspSerLysGluLys-----SerIleLeuLys 278
Db 2416 -----CGAACGATTTAAGAGAACAAAGAACCTTAGAACAAATCATTCATGAA 2466
QY 279 ThrAsnArgGluSerIleThrAsnValGluGlyArg 291
Db 2467 ATGATAAAGAA-----GTTGAAGTCTGT 2490

RESULT 6
US-10-793-626-4160
; Sequence 4160, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OF INVENTION: GRAPHYLOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4160
; LENGTH: 3520
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4160

Alignment Scores:
Pred. No.: 3,286-11 Length: 3520
Score: 235.00 Matches: 89
Percent Similarity: 49.58% Conservative: 86
Best Local Similarity: 25.21% Mismatches: 90
Query Match: 15.04% Indels: 88
DB: 6 Gaps: 17

US-09-837-344-31 (1-316) x US-10-793-626-4160 (1-3520)

QY 4 GlnGlnGlnArgArgAlaLysGluLysLeuGlnGln----- 16
Db 2366 GAACAAAAAACAAAGCT---GAACAAATTAAAGTAAATTTTAAAGTCAAGTCAAG 2422
QY 17 -----GlnSerAspLeuGlnGlnGlnArgLysAlaLysGluLysLeuGlnGln 32
Db 2423 TACAACAATTTAAAGAACAGATACATCAGAAATTTAAAGCTGATACATTAACAAA 2482
QY 33 GlnGlnSerAspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSer 52
Db 2483 CAAGAACACATCTTAAATAATGAACAT-----GAAGAGTTTGAAATTTGAAAAAAT 2533
QY 53 AspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGln 72
Db 2534 GATGAGATATCAAGATGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAT 2593
QY 73 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnSerAspLeuGlnGln---Asp 91
Db 2594 GAG-----ATACAAACAATTTGACCACTAGAAAGTATTTGAAGATATACA 2644
QY 92 ArgLeuAlaLysGlu-----LysLeuGlnGln 101
Db 2645 CAATTATCAAAAGAAAGAAAGCTTCGACATCAAAACAAACAAACAACTACATCAAAA 2704
QY 102 GlnSerAspLeu-----GlnGlnGlnArgArgAlaLysGluLys 114
Db 2705 CAATCTGATTTAGCTGTGTTAAAGACGAAATTAACGAAAGCAAGTTTATGAACGT 2764
QY 115 LeuGlnGlnGlnGlnSerAspLeuGlnGlnGlnArg---ArgAlaLysGluLysLeuGln 133
Db 2765 TTGATTAACAACTTACCGATTCAGAACGTCGAAATTTGAAGTAATGAAAAATCAAA 2824
QY 134 GlnGlnGlnSerAspLeuGlnGlnGlnArgAlaLysGluLysLeuGlnGlnGln 153
Db 2825 TTGTTTAATTCAGATGAATGATGCGTAAAGATGCTTTGAAAGTGAAGAGCAA--- 2881
QY 154 SerAspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnGln 173
Db 2882 -----ATTGACGACAAAGAAATGTAAAGCAAAATTTAAATCAACAATTTAGGAAT 2935
QY 174 GlnGlnGlnArgArg---AlaLysGluLysLeuGlnGlnGlnGlnGlnGlnGln 191
Db 2936 AAACAGCAGCTTAAGATCTTAATGAGAAATGCAATTAATGAACTCAGCTTCAAAA 2995
QY 192 -----GlnGlnArgArgAlaLysGlu--- 198
Db 2996 TGTGATCAGATATACCTTTCTATGAAATCATTCAGATATTAAGCAAAACAAATCA 3055
QY 199 -----LysLeuGln 201
Db 3056 AAGCTAGATGATTAAATCAACCATGCAATGATCATTTAAATGACAGCTATCAACTACA 3115
QY 202 GlnGlnGlnArgAspLeuGlnGlnGlnArgLysAlaAspThrLysLysAsnLeuGlnArgLys 221
Db 3116 GTAGAACGTGCAAGATGCAATGATCTGATGAAACTATTGCAATTTGCCGTAACAAA 3175
QY 222 LysGluLysGluAspLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 241
Db 3176 GTAAATTTAAACGAATGACAAATGATCTGATGAAACTATTGCAATTTGCCGTAACAAA 3232
QY 242 GluLeuProSerGlu---AsnGlnArgGlyTyrTyrIleProHisGlnSerSerLeuPro 260
Db 3233 GAACAATTTGAAGGTTGAATGAACGATACACATTTTAAATGAGCAA----- 3280
QY 261 GlnAspAsnArgLysAsnSerArgAspSerLysGluLys-----SerIleLeuLys 278
Db 3281 -----CGAACGATTTAAGAGAACAAAGAACCTTAGAACAAATCATTCATGAA 3331
QY 279 ThrAsnArgGluSerIleThrAsnValGluGlyArg 291
Db 3332 ATGATAAAGAA-----GTTGAAGTCTGT 3355


```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119
; LENGTH: 8948
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-155-119

Alignment Scores:
Pred. No.: 1,03e-08 Length: 8948
Score: 211.50 Matches: 88
Percent Similarity: 42.18% Conservative: 71
Best Local Similarity: 23.34% Mismatches: 131
Query Match: 13.53% Indels: 87
DB: Gaps: 13

US-09-837-344-31 (1-316) x US-10-623-155-119 (1-8948)

QY 2 AapLeuGluGlnGluArg-----ArgAlaLysGluLysLeuGln 14
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5226 GACCTCCGCGCAGAGGAGCGCTGCATGCCACCTGAGGAAAGCAGAGCCAG 5285

QY 15 GlnGlnGlnSerAapLeu-----GlnGlnAapArgLeuAlaLysGluLysLeuGln 32
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5286 GAAGAGCTGAGAGGCTCTCTCTGAGTCGAGGCCCTGAGCGCGAGTTACTCCAGAA 5345

QY 33 Gln-----GlnSerAapLeuGlnGlnGlnArgAlaLys----- 44
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5346 CAGGAAAGTGTCAAAAGCTCCTTGAGGATGACATTTCCAGAAAGCGGATAGAAAT 5405

QY 45 -----GluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnGlnArgAlaLysGlu 62
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5406 AAAAGCAGAACTTAAATGAAACCAAAATGAAATTTGAGAGGCTGAGCTCTCACAG 5465

QY 63 LysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGln-----ArgAlaLys 78
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5466 AACCTGACCAAGAGAGCATTTGATGTTAGAAAGAACTGCGGAACCTGAGCTGAGTAC 5525

QY 79 GluLysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGlnAapArgLeuAlaLysGluLysLeu 98
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5526 GATGACCTGAGAGAGAGCAGACCAAGCGGACAGTATTAATCA-----ACCATC 5579

QY 99 GlnGlnGlnGlnSerAapLeuGlnGlnGlnGlnGlnArgAlaLysGluLysLeuGlnGln 118
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5580 TTGGAACTAAGAGCCAGCTGCAGATCAGCAACACCGGACCTGGAATCGCAGGGGCTG 5639

QY 119 GlnSerAapLeuGlnGlnGlnGln----- 126
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5640 ATTAATGATTTACAGAGAGAGGAAATTTGAGACAGGAAATTGAGAAATTCAGAAAG 5699

QY 127 -----ArgAlaLysGluLysLeuGlnGlnGlnGlnSerAapLeuGlnGlnGln 143
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5700 CAGGCTTTAGAGGATCTTAATAGGATTCAGGAAATCAAGATCACTGACTCAGGTGCTA 5759

QY 144 ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGlnGlnGln----- 160
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5760 CAGGAAAGAGAGAGCTTCTGCGTGAATAAATCAAGCTCTGAGCAAGCAGAGGCTG 5819

QY 161 ---ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnSerAapLeuGlnGlnGlnGlnGln 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5820 CAGAGGCTGAGAGTGAATCGCAAAATCAACTTACAGGAGGCAAAACAGAGGCTG 5879

QY 180 LysGluLysLeu-----GlnGlnGlnGlnSerAapLeuGlnGln----- 192
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5880 AAACAGGCTGAGAGTGAAGAAACAGCAAAATTCAGAAATCACTGGAATCTGAGGAAACT 5939

QY 193 -----GluArgArgAlaLysGluLysLeuGlnGlnGlnGlnGlnAapLeuGln 208
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5940 CAATATTCCCGCAGAGAGAGGCTATTAGAAAGTAAATTCGAGAAAGAAAGAGAGAG 5999

QY 209 GlnArgLysAlaAapThrLysLysAapLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 228
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6000 AGAGAGAAAGAAAGCTTTAGAGAGTGAAGATCGAAAGACTCCAAGCAGAGATCAAGAAAT 6059
```

```
QY 229 AlaGluAapLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 246
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6060 GAAGAGAGTGCAGAGCGCTTAAGCTGAGATTTCTACAGAGAGACACAGTACAGTTAGAA 6119

QY 247 AapGlnArgGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 266
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6120 ACAGAAACCTCCCGATATCAGAGGAGATGTAATTAATCAAGCAGCGCCCATATAGGCTC 6179

QY 267 SerArgAapSerLys----- 271
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6180 CATCGAGAGACCCAGACTGAGTGTGATGAGACCGTTGACACTCCAGCTGCTTTGAT 6239

QY 272 -----GluLysSerLysLysLysLysLysLysLysLysLysLysLysLysLysLys 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6240 GGGCTGAGGAGAGAGTGCAGATGACAGTACAGTCTATGAGTGTGCTGATCGACAAACA 6299

QY 280 AapArgGluSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 296
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6300 ACCTTGACAAACTA-----TTGAAAGGGAAGAAAGTCACTGGAAGAA 6341

RESULT 11
US-10-821-234-198/C
; Sequence 198, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 198
; LENGTH: 6906
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-198

Alignment Scores:
Pred. No.: 1.17e-08 Length: 6906
Score: 209.50 Matches: 83
Percent Similarity: 46.60% Conservative: 54
Best Local Similarity: 28.23% Mismatches: 106
Query Match: 13.40% Indels: 51
DB: Gaps: 13

US-09-837-344-31 (1-316) x US-10-821-234-198 (1-6906)

QY 3 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 17
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3633 CTGAGAGAGAGACCGGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3574

QY 18 -----SerAapLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 28
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3573 ACTCAGCTCTTAACCTGCTCTCTTCTCCCTGACTGCACAGAGATGAAGCCCGAA 3514

QY 29 Lys-----LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 45
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3513 AAGCTTCCACACTCCAGCGGCGCACAGAAACAGTATTCGGGAG----- 3469

QY 46 LysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3468 ---CTGACAGCTTCAGAGAGAGCGCCGACAGATCGAGCGGAGACTTTCAGATTAACA 3412

QY 66 GluGlnGlnSerAapLeuGlnGln-----GluArgArg 76
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3411 GTTCAGCAAGAGAGACTTCTTCGGGAGAGTCTGTCTCCCGGACCGGTGGAAGCGGAG 3352
```


Oy		72	GlUcInGluInUArg-ArgAlAalYbgIuYvLeuGIInGluGIInGlnSerApLeuGIInAs	91
Db		33020	GAGAGACAGAGACAAGAGAGA CAGAGAGAGGTGGGGAGCTGGAGAGAGAG-AGACAGAG	33078
Oy		91	pArXleuAlAlyBgIuYvLeuGIInGluGIInGlnSerApLeu---GlUcInGluInUArgAr	110
Db		33079	GCAGATGGAGAGACAGAGCATTAAGAGACACAGACAGAGAGTGGGGAGACCGAGAGAGA	33138
Oy		110	gAlAlyBgIuYvLeuGIInGluGIInGlnSerApLeuGIInGluGIInUArgAlAlyBgI	130
Db		33139	GAG-AGACAGAGCCAGATGGAGAGACAGAGACAGAGAGACAGAGAGAGCGGGGAGATG	33197
Oy		130	uYvLeuGIInGluGIInGlnSerApLeuGIInGluInUArg-----	143
Db		33198	GAGAGAGAGACAGAGCAGAGGCCAGATGGAGAAACAGAGATTAGAGACAGACAGAGAGTG	33257
Oy		144	----ArgAlAlyBgIuYvLeuGIInGluGIInGlnSerApLeuGIInGluInUArg-----	160
Db		33258	CGGGAGATTGAGAGACAGAGACAGAGCCAGATGGAGAGACAGAAAAGACAGACAGAG	33317
Oy		161	-----	172
Db		33318	AGCTGGCGGAGATTGAGAGACAGAGCACAGAGCACAGCCAGATGGAGAGACAGAGATTAGAGA	33377
Oy		172	pLeuGIInGluInUArgAlAlyBgIuYvLeuGIInGluGIInGlnSerApLeuGIInGlu	192
Db		33378	GACAGAGAGAGAGAGGTGGGGGAGACGGAGAGAGAGAGACAGAGCCAGATTGGAGAGAC	33437
Oy		192	nGUInUArgAlAlyBgIuYvLeuGIInGluGIInUArgApLeuGIInUAryBwAl	212
Db		33438	AGAGACAGAGAGACAGAGAGAGCGCGGGGAGATTGAGAGAGAGACAGAG---ACAGAGCC	33494
Oy		212	AABPThLyvLYvAhvLnLcuGLrGLvLYvGLuHhVgLIyAPhILEuAlAGLuAspLE-	232
Db		33495	AGATGGAGAGACAGAGACAGACAGACAGACAGAGAG-----	33528
Oy		232	uTYrGLiArXpLeuGIuLlEProAlAlEGluLeuProserGLuAnGluInUAryTYrry	252
Db		33529	-GTGGGGAGGTGGAGAGACAGACAGAGCCAGATGGAGAGACAGAGATTAGAGAGAGA	33587
Oy		252	rIEPrOhVgInSerSertLeuProGlnApeAmAryGLiYasSerArYAsPserLyBgI	272
Db		33588	GACAGAGAGCGGGGAGATGGAGAGACAGAGACAGAGCGCATGTGAGAGACAGAGAT--	33645
Oy		272	uIlEsErILlEGluYvThrAmAryGluSerILeThThAsvAlGluClYAryAr	292
Db		33646	-----AGAGAGACAGACAGAGAGTGGGGGAGATGGAGAGACAGAGACAGAG	33692
Oy		292	gAPhILhIEtLYvGLiYvLeuGIInGluYvLYvAPhGLISerILEyProGluInLy	312
Db		33693	GCAGATGGAGAGACAGAGATTAGAGACACAGACAGAGAGTGGGGGAGATTGAGAGAGAGA	33752
Oy		312	gLUApLySeR 316	
Db		33753	CAGAGACAGAGAC 33765	
RESULT 15				
US-10-821-234-49				
; Sequence 49, Application US/10821234				
; Publication No. US20050255114A1				
; GENERAL INFORMATION:				
; APPLICANT: Labat, Ivan				
; APPLICANT: Strache-Crain, Birgitte				
; APPLICANT: Andarmant, Susan				
; APPLICANT: Tang, Y. Tom				
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia				
; FILE REFERENCE: 821A				
; CURRENT APPLICATION NUMBER: US/10/821,234				
; CURRENT FILING DATE: 2004-04-07				
; PRIOR APPLICATION NUMBER: US 60/462,047				
; PRIOR FILING DATE: 2003-04-07				
; NUMBER OF SEQ ID NOS: 1704				

```

: SOFTWARE:pt_seq_genes Version 1.0
: SEQ ID NO 49
: LENGTH: 5468
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-821-234-49

Alignment Scores:
Pred. No.: 1,14e-07 Length: 5468
Score: 197.00 Matches: 91
Percent Similarity: 43.82% Conservative: 72
Best Local Similarity: 24.46% Mismatches: 139
Query Match: 12.60% Indels: 71
DB: 6 Gaps: 16

US-09-837-344-31 (1-316) x US-10-821-234-49 (1-5468)
QY 1 SerAspLeuGlnGlnGlnArgArgAlaLys-----GluLysLeuGlnGlnGlnSerAsp 19
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3174 AGCCAGGTGGAGACAGACAGAAAGCAGAGAGCTGGCCAAAGCTTCGACAGAGCTCAGCAG 3233
QY 20 LeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln 39
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3234 GTCACAAAGAGCTGTGGAGAACTCAGAGCTGTGGCAGAAAT-----GACACAG 3284
QY 40 GluArgArgAla-----LysGluLysLeuGlnGln 50
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3285 CAGGCGAAAGCTTCGAAACCAAGGACGAGCTGCTTCGAGAAACAGAGTCTCGACGCTGACG 3344
QY 51 GlnSerAspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGln----- 66
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3345 GCGTCCACAGAGAGAGTGAAGAGAGCCCTGCGAAGAGCGCTGCGACGAGTCAAGCCGGAG 3404
QY 67 -----GlnGlnSerAspLeuGlnGlnArgArgAlaLysGluLysLeuGlnGln 83
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3405 CTGTGCCACACGACAGACAGCCAGCCAGCTCTCGGCGGAGATGCGGAGAGGCCACAGAG 3464
QY 84 GlnGlnSerAspLeuGlnGlnAspArgLeuAlaLysGluLysLeuGlnGlnGlnSer 103
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3465 CAACACGACGACGATGGCCGAG-----CTGCACACGAAATTCACAGTCTCCGAGGCG 3515
QY 104 AspLeuGlnGlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln 123
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3516 GAGGNG-----CCAGCAAAATGCGAGAGAGCTGATGGCTCCACGGGACACTCCAG 3566
QY 124 GlnGlnArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln----- 141
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3567 GAGGCGAGGGGGGAGAACTCCCACTTACACAGAAATCCGATTCATTGAGGCTTGCTG 3626
QY 142 -----GluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGln 158
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3627 GAGGCGGGCCACGGCGGGAGATGCCACAGAGACTCCAGAGCCACGACGGGGAGGCTCAG 3686
QY 159 GluArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGln----- 176
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3687 CAGCAGACTCGCTCCAGAGAGACTGGAAGTCCAGGTGTGGGTGTGGAGAGAGAGGCATC 3746
QY 177 -----ArgArgAlaLysGluLysLeuGlnGlnGlnGlnSerAspLeuGlnGln 194
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3747 GAGCTCAGGAGGAGCTGTCCAGACACAGAAAGTGAAGAAACATGCTCCGGGAGAGAAC 3806
QY 195 -----ArgAlaLysGluLysLeuGlnGlnGlnGlnGlnArgAsp 206
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3807 TGGAAAGCCATCGAGAGCACTGGCCACAGCGCCGAGAGGCGCTCGAAGAGAAAGCTCACCTC 3866
QY 207 LeuGlnGlnArgLysAlaAspThrLysLysAsnLeuGlnArgLys-----LysGluHis 224
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3867 CTGACCCGACGAGCGAAGAGAAATCGAAGAACGACTCTG-TCTGATTGAGGCGCAGACAT 3925
QY 225 GlyAspIleLeuAlaGluAsp-----LeuThrGlyArgLeuGlnIleProAla 240
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 3926 GAGAGCCCTGCTGCTGCTCTCCAGAACTCTGCTGCTTGGCACACAGAAATTCACACCGA 3985

```

```
OY      241  ILeGIuLeuProSeRGluAaNGluArgGlyTyTyTyleProHisGlnSerSerLeuPro 260
      DB      3986  GTGGCTGCAGCATCTCAAGAGAAAG-----CCCCACGCTGCTGAGCACCCC 4033
      OY      261  GlnAspAsnArgGly-AsnSerArgAsp--SerLySGluIleSerIleIleGluLysTh 279
      DB      4034  GCCAGCTCCCGGAGAGCCCTCTCGGACCTGACCTCCAGTTGAGGGAGGCCGAGGAGAC 4093
      OY      279  rAsnArg-----GluSerIleThrThrAsnValGluG 290
      DB      4094  GCAGAGCACACTGCAGGCCGAGTGTGACCACTACCGCAGCATCTGCGGAGACGAGGG 4153
      OY      290  YArg--ArgAspIleHisLySGlyHisLeuGluGluLys-----LysAs 304
      DB      4154  CATGCTCAGAGACCTGCAGAAAGCCTGAGAGAGAGAGACAGAGCTGTGAGGGCCAAAGT 4213
      OY      304  pGIySerIleLysProGluGluLysGluAspLys 315
      DB      4214  GGGCGCGCGAGAGAGAGACTCCAGAAATCCCGG 4247
```

Search completed: December 7, 2005, 16:25:08
Job time : 443.478 secs

Db 1421 SDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRA 1480
Qy 61 KEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 120
Db 1481 KEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 1540
Qy 121 DLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKE 180
Db 1541 DLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKE 1600
Qy 181 EKLQEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLE 240
Db 1601 EKLQEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLE 1660
Qy 241 TELPENERGYIIPHOSSLPONRGNRSRKEISITTEKTNRESITTNVEGRDTHKGLH 300
Db 1661 TELPENERGYIIPHOSSLPONRGNRSRKEISITTEKTNRESITTNVEGRDTHKGLH 1720
Qy 301 EKKDSIKPEQEKDS 316
Db 1721 EKKDSIKPEQEKDS 1736

RESULT 2

S28589
trichohyalin - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 12-Mar-1993 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: S28589
R:Fliez, M.J.; Rogers, G.E.
submitted to the EMBL Data Library, December 1992
A:Description: Examination of the gene encoding rabbit trichohyalin.
A:Reference number: S28589
A:Accession: S28589
A:Molecule type: DNA
A:Residues: 1-1407 <F1E>
A:Cross-references: UNIPROT:P37709; UNIPARC:UPI000013738C; EMBL:Z19092; NID:g1746; PIDN:
C:Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root sh
C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
C:GeneID:8
A:introns: 46/3
A:Superfamily: trichohyalin; calmodulin repeat homology
C:Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
P:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 21.0%; Score 328; DB 1; Length 1407;
Best Local Similarity 29.4%; Pred. No. 3e-08;
Matches 88; Conservative 77; Mismatches 92; Indels 42; Gaps 10;
Qy 4 EBERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 54
Db 150 EBERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 208
Qy 55 EBERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 113
Db 209 EBERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 264
Qy 114 KLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKE 166
Db 265 EBERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 323
Qy 167 QEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKE 226
Db 324 EBERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 370
Qy 227 ILAEDLYGRLEIPATELPENERGYIIP--HOSSLPONRGNRSRKEISITTEKTNRESITTNVEGRDTH 282
Db 371 ILAEDLYGRLEIPATELPENERGYIIP--HOSSLPONRGNRSRKEISITTEKTNRESITTNVEGRDTH 426

RESULT 3
T15597
hypothetical protein C25A11.4b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15597
R:Favell, T.
submitted to the EMBL Data Library, October 1995
A:Description: The sequence of C. elegans cosmid C25A11.
A:Reference number: 218375
A:Accession: T15597
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-737 <FAV>
A:Cross-references: UNIPARC:UPI00001787BA; EMBL:U39650; NID:g1049376; PID:g1049381; PIDN:
A:Experimental source: strain Bristol N2
C:GeneID:8
A:introns: 65/3; 88/3; 126/3; 433/2; 494/2; 711/2

Query Match 20.4%; Score 318.5; DB 2; Length 737;
Best Local Similarity 26.9%; Pred. No. 4.6e-08;
Matches 90; Conservative 91; Mismatches 104; Indels 49; Gaps 12;

Qy 3 LEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 62
Db 237 LEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 292
Qy 63 KLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKE 121
Db 293 RIETRIIRIGRIERIRERRERREKKAEDLLERLELER--IERERRELEARRQOLE 350
Qy 122 L--EQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 162
Db 351 LQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLE 410
Qy 163 KEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKE 211
Db 411 RIETRIIRIGRIERIRERRERREKKAEDLLERLELER--IERERRELEARRQOLE 469
Qy 212 ADTKNLRRKKEHGDILAEEDLYGRLEIPATELPENERGYIIPHOSSLPONRGN 266
Db 470 EBERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 529
Qy 267 SRDSKEISITTEKTNRESITTNVEGRDTHKGLH 300
Db 530 ELDRAKQELSEREME-----KERRDRERANEE 557

RESULT 4

T15598
hypothetical protein C25A11.4a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15598
R:Favell, T.
submitted to the EMBL Data Library, October 1995
A:Description: The sequence of C. elegans cosmid C25A11.
A:Reference number: 218375
A:Accession: T15598
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1017 <FAV>
A:Cross-references: UNIPARC:UPI00001787BA; EMBL:U39650; NID:g1049376; PID:g1049380; PIDN:
A:Experimental source: strain Bristol N2
C:GeneID:8
A:introns: 65/3; 88/3; 126/3; 433/2; 494/2; 711/2; 735/1; 792/3; 833/3; 873/2

Query Match 20.4%; Score 318.5; DB 2; Length 1017;
Best Local Similarity 26.9%; Pred. No. 6.1e-08;
Matches 90; Conservative 91; Mismatches 104; Indels 49; Gaps 12;

Qy 3 LEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 62
Db 237 LEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOSDLEQERRAKEKLOEQOS 292


```

M2 protein precursor - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
A:Variatey: serotype M2
C:Date: 22-Nov-1993 #sequence revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S23325; S35761; S61078; S60785
R:Bessen, D.E.; Flechtel, V.A.
Intect. Immun. 60, 124-135, 1992
A:title: Nucleotide sequences of two adjacent M or M-like protein genes of group A strept.
A:Reference number: S23325; PMID:92104662; PMID:1370269
A:Accession: S23325
A:Molecule type: preliminary
A:Status: preliminary
A:Residues: 1-407 <BES>
R:Podbielski, A.
submitted to the EMBL Data Library, November 1992
A:Reference number: S35760
A:Accession: S35761
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-17 <POD>
A:Cross-references: UNIPARC:UPI000017823B; EMBL:X69324
R:Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kenoe, M.
submitted to the EMBL Data Library, July 1994
A:Description: Noncongruent relationships between variation in emm1 gene sequences and t
A:Reference number: S61072
A:Accession: S61078
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 12-94 <MHA>
A:Cross-references: UNIPARC:UPI0000081BF6; EMBL:U11958; NID:9535393; PIDD:AAA9574.1; PI
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kenoe, M.A.
Mol. Microbiol. 14, 619-631, 1994
A:title: Non-congruent relationships between variation in emm gene sequences and the pop
A:Reference number: S60784; PMID:95198537; PMID:7891551
A:Accession: S60785
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 31-89 <MH2>
A:Cross-references: UNIPARC:UPI000017823C; EMBL:U11958
C:Superfamily: M5 protein

Query Match          18.5%; Score 289; DB 2; Length 407;
Best Local Similarity 29.6%; Pred. No. 6.2e-07;
Matches    89; Conservative   69; Mismatches 105; Indels   38; Gaps   12;

QY      4 EERRAKETLQEQOSDLQ-----DLRAKETLQEQOSDLQERRAKEKLEQEOSDLQE 57
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB     90 EEHKDKHELTKEKSEVENHYLRDLQDYKEG-QERQKNLE-----LERQSREVE 140
QY     58 RRAKETLQEQOSDLQERRAKEKLEQ--QSDLEORLAKETLQEQOSDLQERRAKEKLQ 116
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    141 KYEQEQLQQQ-QLEKEKISFASRKSLRDLEASRAKKDLFAHQTKLEBKQISEASR 199
QY    117 EQ-QSDDLEERRAKEKLEQOSDLQERRAKEKLEQOS-DLEERRAKEKLEQOSDLE 174
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    200 KSLRLDEASRAKKDLFAHQTKLEEKQISEASRQGLSRDLEASRAKKDLFAHQTKL 259
QY    175 QERRAKEKLEQOS-DLEERRAKEKLEQOSDLQ-----RKADTKNLRKKENGDILAE- 230
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    260 EEKQISEASRQGLSRDLEASREAKKV---EADLEANSKIOLAKLNKLEEGKLSLEK 316
QY    231 ---DIAGRIEIPAIELPSE-----NERGYIPIHQSLLPQDRNGNSRDSKEISILEK 278
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    317 EKAELOAKLEAKKALKEQLAKQAEELATLKGNQTPNAVAQAARRSAMTQQRITPS 376
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      279 T 279
       |
DB     377 T 377

```

trichhyalin like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: B85431
R:Anonymous, The European Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; WUID:20083488; PMID:10617198
A/Accession: B85431
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1432 <STO>
A/Cross-references: UNIPROT:O23230; UNIPARC:UPI00000092F0; GB:NC_001268; NID:g7270600; P
C:Genetics:
A:Gene: AT4G16520
A/Map position: 4

Query Match 18.4%; Score 287.5; DB 2; Length 1432;
Best Local Similarity 27.9%; Pred. No. 2.1e-06;
Matches 103; Conservative 73; Mismatches 92; Indels 101; Gaps 20;

Qy 3 LQERRAKKQLQEQSDLEQDLAKKQLQEQ--QSDLEQERRAKKQLQEQSDLEQERR 59
Db MEETIKKRLRLREND-RRRVAVKAKENKRLKALQEQEK-ERIKKAKAKENRR 689
Qy 60 AKK-----KLQEQSDLEQERRAKKQLQEQSD-----LEQDLAKKQLQEQ 102
Db AVEARKAKQERRKKEQ--ELEQLKAFKFEENRRMRKAPALQEK--ERRIKAR 744
Qy 103 SLEQERR--AKKQLQEQ--SDLEQERRAKKQLQEQSDLEQERRAKKQLQEQSDL 156
Db 745 EKENERRIKKAKKQLQEQRLKATLEQEKERO-IKEROEENERRAKVLEQ---A 799
Qy 157 EQERRAKKQLQEQSDLEQERRAKK-----KLQEQSDL-----EQER 194
Db 800 ENRRKLEKLEQK-----ENRRKLEKTEKKEENKKLAEALBLEKERRLIAFERAER 855
Qy 195 RAKKQLQEQORDLEQRKATYKQLERK-KEHGDTLAEDLYGRLEIPALEPSENERGYI 253
Db 856 RLKEDLEQEMRMRLQAKERRRLHRENOH-----QENRK--- 892
Qy 254 PHQSSLPQDNKGRSPDKESIIEKTRRESTTTNVEGRDTHKG--HLKKK---DGS 306
Db 893 QHRYs-----GESDEKERDACEKTCETTKAHGQSSNESLDTLEENESIDNDVS 946
Qy 307 IKPEQEKDK 315
Db 947 VKKKKEE 955

RESULT 9
G70241
hyoethetical protein BB116 - Lyme disease spirochete plasmid I/1p28-4
C/Species: Borrelia burgdorferi (lyme disease spirochete)
C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C/Accession: G70241
R:Fraser, C.M.; Casjans, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit-
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A/Reference number: A70100; WUID:98065943; PMID:9403685
A/Accession: G70241
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-451 <KLE>
A/Cross-references: UNIPROT:O50870; UNIPARC:UPI00000568B1; GB:AE000789; NID:g2690079; P
A/Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 18.1%; Score 283.5; DB 2; Length 451;

Best Local Similarity 28.2%; Pred. No. 1.2e-06;
Matches 97; Conservative 74; Mismatches 122; Indels 51; Gaps 18;

QY 1 SDLEQERRAKEKLOEQSDLEQDRLAKEKLOEQ-SDLEQERRAKEKLOEQ-SDLEQER 58
D 40 TSSKQKSPKREBEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEEL 99
QY 59 RAKEKLOEQ-SDLEQERRAKEKLOEQ-SDLEQDRLAKEKLOEQSDLEQERRAKEK 115
D 100 ELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQ 159
QY 116 QEOQ-SDLEQERRAKEKLOEQ-SDLEQERRAKEKLOEQ-SDLEQERRAKEKLOEQ-S 171
D 160 QOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEEL 219
QY 172 DLEQERRAKEKLOEQSDLEQERRAKEKLOEQ-RDLEQ-R-KADTKK 218
D 220 KKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQOEELKKQ 279
QY 219 ERKKEHODILAEDYGLER-PALPSEN-ERGY-YIPHOSSL-PODNRGNSR 268
D 280 EKE-PEDHYGNTSFRGLNMGPTEDISNTERSTIRYRHYTVLSPD-PH 328
QY 269 DSKEI-SIEKTNESITTN-VEGRDIIHKGHLEBKD 304
D 329 ELKEFANIOTDKLAASVAFNSFSAIGALDVLSDHLYPKD 372

RESULT 10

E84565
hypothetical protein At2g18540 (imported) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: E84565
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.;
Euse, D.; Niekman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402: 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: E84565
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-699 <STO>
A/Cross-references: UNIPROT:Q9ZU69; UNIPARC:UPI00000A5B25; GB:A8002093; NID:q4218005; PI
C/Genetics:
A/Gene: At2g18540
A/Map position: 2

Query Match 17.4%; Score 272; DB 2; Length 699;

Best Local Similarity 28.9%; Pred. No. 5.8e-06;
Matches 74; Conservative 69; Mismatches 87; Indels 26; Gaps 9;

QY 4 EOEERRAKEKLOEQSDLEQDRLAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRA 60
D 426 ERKRRESEETERRKEEBAKREBAKREBESEKREBESEKREBESEKREBESEKREBE 485
QY 61 KEKLOEQSDLEQERRAKEKLOEQSDLEQDRLAKEKLOEQSDLEQERRAKEKLOEQ-EOQ 119
D 486 RE-EEBAKREBEKREBEBAEQARKREER-EKEEBAKREBEKREBEKREBEKREBE 540
QY 120 SDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQ 173
D 541 RREQERKRR-EEARKREBEKREBEBAKRR-EOERQKREBEKREBEKREBEKREBE 594
QY 174 EOEERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKE 233
D 595 REEBAKREBEKREBEKREBEKREBEKREBEKREBEKREBEKREBEKREBEKREBE 650
QY 234 GRLEIPALPSENER 249
D 651 -RREBAKREBEKREBEKREBEKREBEKREBEKREBEKREBEKREBEKREBEKREBE 650

RESULT 11

G88484
protein F23F12.8 (imported) - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: G88484
R/Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.e
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A/Accession: G88484
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-887 <STO>
A/Cross-references: UNIPROT:P46504; UNIPARC:UPI0000060FC9; GB:chr_III; PIDN:AAA20610.1;
A/Note: exon 5 similar to trichoyalin and tropomyosin
C/Genetics:
A/Gene: F23F12.8
A/Map position: 3

Query Match 17.4%; Score 271.5; DB 2; Length 887;
Best Local Similarity 25.0%; Pred. No. 7.5e-06;
Matches 83; Conservative 93; Mismatches 105; Indels 51; Gaps 13;

QY 6 ERRAKEKLOEQSDLEQDRLAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQ 57
D 310 EROQOEKFEK-MEOERLROEK-EKAKELERRRLESESTAPARQELDROATYAE 364
QY 58 RRAKEKLOEQSDLEQERRAKEKLOEQSDLEQDRLAKEKLOEQSDLEQERRAKEKLOEQSDLEQ 112
D 365 RMMERRELERLERLEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEK 419
QY 113 EKLQEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSD 172
D 420 QELBAARKYQLQEBEQRKIQQKQVMEQIRQOEBAKQOELRVLEEE-RARELERVQEE 478
QY 173 LEOERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSD 224
D 479 LERQHQ-MEILROQOEBOKKLEKREKREKREKREKREKREKREKREKREKREKREKREK 537
QY 225 GDILAEDYGLERLIPALPSENERGYIIPHOSSLPODNRGNSRDSKEISIEKTN-RES 283
D 538 RKULEKEM-EDRONAIYEEERRILKEER-RKQIEERRIQOQ 580
QY 284 ITTNVEGRDIIHKGHLEKKDGSIKPEKEDK 315
D 581 IMATERSRLDAMERERKRLQIKSEKQK 612

RESULT 12

S37046
IGA receptor - Streptococcus pyogenes
C/Species: Streptococcus pyogenes
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: S37046
R/Heden, L.; Lindahl, G.
submitted to the EMBL Data Library, May 1993
A/Description: Conserved and variable regions in protein Arp, the IGA receptor of Strept
A/Reference number: S37046
A/Accession: S37046
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-402 <HED>
A/Cross-references: UNIPROT:Q54876; UNIPARC:UPI00000BC74D; EMBL:Z22751; NID:G397628; PI
C/Superfamily: MS protein
C/Keywords: Immunoglobulin receptor

Query Match 17.2%; Score 269; DB 2; Length 402;
Best Local Similarity 27.9%; Pred. No. 5e-06;
Matches 93; Conservative 63; Mismatches 105; Indels 72; Gaps 12;

A; Introns: 622/3

Query Match

17.1%; Score 267.5; DB 2; Length 729;

Best Local Similarity 27.2%; Pred. No. 9.6e-06;

Matches 93; Conservative 70; Mismatches 110; Indels 69; Gaps 13;

```
QY 2 DLBERRAKELQEQOSDLEODRLAKELQEQOSDLEERRAKELQEQOSDLEERRAK 61
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 285 DLIRKKEERLQEQNH--EHDQLYQAVQ-----LDIMRKKEESRERDQKYEDELAV 337
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 EKLO-----EQOSDLEERRAKELQEQOSDLEQDR---LAKELQEQOSDLEQER 109
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 338 SKPQLDLVKRKEERARHLEERK--KEAFQEQQLAQAYELDMIKRREERREEREEHR 396
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 110 RAKE-----KLQEQOSDLEERRAKELQEQOSDLEQER--RAKELQEQOSDLEQERRA 162
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 397 REMEYEQOHQHLBAKFNHDMIRQEEERRRAAAREEEKILIRAKERLERLEREEERERKQ 456
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 163 KE-----KLQEQOSDLEERRAKELQEQO-----SDLEERRAKE-----KLQEQOR 205
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 457 KLEEELELRKAKEELDMRREKKEERERLRREYELKLEERKAEERKQEEEREA 516
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 206 DLEQRKADTKKGLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSIPODNRG 265
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 517 AVARWKAKAEARLAKEQEEREM-----EREY-----QRRLQEDLIN 553
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 266 NSRDSKEISIIKTNRESITTVNQRDINHGLEKKDGI 307
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 554 SGLDEKAIIAILK--KEKIKKEKKGSDSDSEKKAQKQGI 593
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 15

150591

class II INCENP protein - chicken

C:Species: Gallus gallus (chicken)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C:Accession: 150591

R:McKAY, A.M.; Eckley, D.M.; Chue, C.; Earnshaw, W.C.

J. Cell Biol. 123, 373-385, 1993

A:Title: Molecular analysis of the INCENP (inner centromere proteins): separate domains

A:Reference number: A48482; MUID:94012983; PMID:8408220

A:Accession: 150591

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-877 <MAC>

A:Cross-references: UNIPROT:P5352; UNIPARC:UPI000012D654; EMBL:Z25420; NID:g414110; PIR

Query Match 17.0%; Score 266; DB 2; Length 877;

Best Local Similarity 28.1%; Pred. No. 1.3e-05;

Matches 73; Conservative 74; Mismatches 77; Indels 36; Gaps 9;

```
QY 4 EQERRAKELQEQOSDLEODRLAKELQEQOSDLEERRAKELQEQOSDLEERRAK 63
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 501 DPKEKERQQLAKKEAEALQKQKVEEKKRQEE--AKLRREERLRYQLARERAEQ 558
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 64 LQEQOSDLEERRAKELQEQOSDLEODRLAKELQEQOSDLEERRAKELQEQOS 120
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 559 LEE-----ERKRRIEQKLALFDEKTEKAREERLAEKI-----KGRAAAKQWEAEA 605
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 121 DLEERRAKELQEQOSDLEERRAKELQEQOSDLEERRAKELQEQOSDLEQER--- 177
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 606 RRRDDEERAKQALQOE--EEERHKEIMQKKE--EEQERARKIAEQROAEQERKQLA 661
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 178 --RAKELQEQOSDLEERRAKELQEQO-----QRDLEQRKADTKKGLERKKEHGDIL 228
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 662 AEREQERKKEQERKKEERIRQAEKQEQEKAARLQKEVLAKEQLQKEMKEKEEQLL 721
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 229 AEDLYGRLEIPALPSENE 248
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 722 AE--MKRQEQEQKLLPREEQK 739
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

Search completed: December 7, 2005, 07:21:24
Job time : 25.0831 secs

THIS PAGE BLANK (USPTO)


```
DB 245 IELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRESITTNVEGRDIIKHGHL 304
QY 301 EKKDGSIKPEQKEDKS 316
DB 305 EKKDGSIKPEQKEDKS 320

RESULT 2
Q25893 PLAF7
ID Q25893 PLAF7 PRELIMINARY; PRT; 1909 AA.
AC Q25893;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Liver stage antigen.
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NF54;
RX MEDLINE=92107224; PubMed=1840628; DOI=10.1016/0166-6851(91)90117-0;
RA Zhu J., Hollingdale M.R.;
RT "Structure of Plasmodium falciparum liver stage antigen-1."
RL Moll. Biochem. Parasitol. 48:223-226(1991).
DR EMBL; X56203; CAA39663.1; -; Genomic_DNA.
DR PIR; S24597; A45592.
DR HSSP; O15813; 1D7M.
SQ SEQUENCE 1909 AA; 230153 MW; 8B61A524C797217 CRC64;

Query Match 95.6%; Score 1494; DB 2; Length 1909;
Best Local Similarity 95.9%; Pred. No. 2.4e-50;
Matches 303; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 SDLEORRAKEKLEQOQSDLEODRLAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 60
DB 1421 SDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 1480
QY 61 KEKLEQOQSDLEORRAKEKLEQOQSDLEODRLAKEKLEQOQSDLEORRAKEKLEQOQ 120
DB 1481 KEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQ 1540
QY 121 DLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 180
DB 1541 DLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 1600
QY 181 EKLQEQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLE 240
DB 1601 EKLQEQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLE 1660
QY 241 IELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRESITTNVEGRDIIKHGHL 300
DB 1661 IELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRESITTNVEGRDIIKHGHL 1720
QY 301 EKKDGSIKPEQKEDKS 316
DB 1721 EKKDGSIKPEQKEDKS 1736

RESULT 3
Q25893 PLAF7
ID Q25893 PLAF7 PRELIMINARY; PRT; 1596 AA.
AC Q25893;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Liver stage antigen, putative.
GN ORFNames=PF10_0356;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=36329;
```

```
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiolini S.,
RA Perera M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carnocci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Frazer C.W., Barrett B.G.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014834; AAN35553.1; -; Genomic DNA.
SQ SEQUENCE 1596 AA; 191691 MW; C0B3A2C0F5D60AC CRC64;

Query Match 93.0%; Score 1453.5; DB 2; Length 1596;
Best Local Similarity 90.1%; Pred. No. 7.7e-49;
Matches 300; Conservative 5; Mismatches 11; Indels 17; Gaps 1;

QY 1 SDLEORRAKEKLEQOQSDLEODRLAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 60
DB 1091 SDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 1150
QY 61 KEKLEQOQSDLEORRAKEKLEQOQSDLEODRLAKEKLEQOQSDLEORRAKEKLEQOQ 120
DB 1151 KEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQ 1210
QY 121 DLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 163
DB 1211 DLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRA 1270
QY 164 EKLQEQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLE 223
DB 1271 EKLQEQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLEORRAKEKLEQOQSDLE 1330
QY 224 HGDVLAEDLYGRLPALELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRES 283
DB 1331 HGDVLAEDLYGRLPALELPSENERGYIIPHOSSLPODNRGSRDSKEISITTEKTNRES 1390
QY 284 ITTNVEGRDIIKHGLEEKDGSIKPEQKEDKS 316
DB 1391 ITTNVEGRDIIKHGLEEKDGSIKPEQKEDKS 1423

RESULT 4
Q25887 PLAF7
ID Q25887 PLAF7 PRELIMINARY; PRT; 318 AA.
AC Q25887;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Liver stage antigen-1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94267224; PubMed=7515922;
RA Fidock D.A., Grais-Rasse H., Lepeux J., Brahim K., Benmohamed L.,
RA Wellouk S., Guerin-Marchand C., Londono A., Raharimalala L.,
RA Meis J.F., Langsley G., Roussillon C., Tartar A., Druilhe P.;
RT "Plasmodium falciparum liver stage antigen-1 is well conserved and
RT contains potent B and T cell determinants."
RL J. Immunol. 153:190-204(1994).
DR EMBL; Z50319; CAA82974.1; -; Genomic_DNA.
DR HSSP; O15813; 1D7M.
FT NON TER 318
SQ SEQUENCE 318 AA; 38098 MW; 989F32A70605DFB2 CRC64;
```

```
Query Match      49.4%; Score 771.5; DB 2; Length 318;
Best Local Similarity 86.2%; Pred. No. 5.3e-23;
Matches 163; Conservative 12; Mismatches 13; Indels 1; Gaps 1;

Qy 2 DLEOERAKKXKLEQOQSDLEODRLAKEKLEQOQSDLEOERAKKXKLEQOQSDLEOERAKK 61
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 131 DDDKKXIKKODENRQEDLE-EKAKAKKLEQOQSDSDEERAKKXKLEQOQSDLEQERLAK 169

Qy 62 EKLQEQQSDLEOERAKKXKLEQOQSDLEODRLAKEKLEQOQSDLEOERAKKXKLEQOQSD 121
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 190 EKLQEQQSDLEOERAKKXKLEQOQSDLEOERLAKKXKLEQOQSDLEOERAKKXKLEQOQSD 249

Qy 122 LEOERAKKXKLEQOQSDLEOERAKKXKLEQOQSDLEOERAKKXKLEQOQSDLEOERAKE 181
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 250 LEOERRAKKXKLEQOQSDLEOERLAKKXKLEQOQSDLEODRLAKEKLEQOQSDLEOERAKE 309

Qy 182 KLEQEQQSDL 190
    ||| : : : : : |||
Db 310 RLEQEQQSDL 318

RESULT 5
Q25844_PLAFA PRELIMINARY; PRT; 280 AA.
ID Q25844_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25844;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40885; AAB59231.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32940 MW; E9708B3CFPAE9CF CRC64;

Query Match      35.4%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 RKADTKNLERKKEHGDIILADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 269
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 1 RKADTKNLERKKEHGDIILADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60

Qy 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 316
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 6
Q25849_PLAFA PRELIMINARY; PRT; 280 AA.
ID Q25849_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25849;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40890; AAC41600.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEB9CF CRC64;

Query Match      35.4%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 RKADTKNLERKKEHGDIILADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 269
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 1 RKADTKNLERKKEHGDIILADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60

Qy 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 316
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 7
Q25850_PLAFA PRELIMINARY; PRT; 280 AA.
ID Q25850_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25850;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40891; AAC41601.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 280 AA; 32989 MW; E96F812CFPABED8 CRC64;

Query Match      35.4%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 RKADTKNLERKKEHGDIILADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 269
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 1 RKADTKNLERKKEHGDIILADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60

Qy 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 316
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 8
Q25843_PLAFA PRELIMINARY; PRT; 280 AA.
ID Q25843_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25843;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
```

```

OX  NCBI_TaxID=5833;
RN  NUCLEOTIDE SEQUENCE.
RP  MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RT  "Sequence variations in the non-repetitive regions of the liver stage-
RT  specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT  isolates."
RL  Mol. Biochem. Parasitol. 71:291-294(1995).
DR  EMBL, L40884; AAB59230.1; -; Genomic_DNA.
FT  NON TER
SQ  SEQUENCE 280 AA; 32926 MW; 467080932FAAD33 CRC64;

Query Match      35.4%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.3e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  210 RKAATKKKLERKKEHGDILAEADLYGRLEIPALTELPSENERGYIIPHOSLPODNRGNSRD 269
    |||||
DB  1 RKAATKKKLERKKEHGDVLAEDLYGRLEIPALTELPSENERGYIIPHOSLPODNRGNSRD 60

OY  270 SKEISIIKTNRESITTNVEGRDRIHKHGLEKKDGSIKPEQKEDKS 316
    |||||
DB  61 SKEISIIKTNRESITTNVEGRDRIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 9
O25846 PLAF A
ID  O25846 PLAF A PRELIMINARY; PRT; 280 AA.
AC  O25846;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Liver stage-specific antigen 1 (Fragment).
GN  Name=LSA-1;
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  NUCLEOTIDE SEQUENCE.
RP  MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RT  "Sequence variations in the non-repetitive regions of the liver stage-
RT  specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT  isolates."
RL  Mol. Biochem. Parasitol. 71:291-294(1995).
DR  EMBL, L40887; AAC41597.1; -; Genomic_DNA.
FT  NON TER
SQ  SEQUENCE 280 AA; 32882 MW; 46763641C81AFC33 CRC64;

Query Match      35.4%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.3e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  210 RKAATKKKLERKKEHGDILAEADLYGRLEIPALTELPSENERGYIIPHOSLPODNRGNSRD 269
    |||||
DB  1 RKAATKKKLERKKEHGDVLAEDLYGRLEIPALTELPSENERGYIIPHOSLPODNRGNSRD 60

OY  270 SKEISIIKTNRESITTNVEGRDRIHKHGLEKKDGSIKPEQKEDKS 316
    |||||
DB  61 SKEISIIKTNRESITTNVEGRDRIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 10
O25847 PLAF A
ID  O25847 PLAF A PRELIMINARY; PRT; 280 AA.
AC  O25847;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Liver stage-specific antigen 1 (Fragment).

```

```

GN  Name=LSA-1;
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  NUCLEOTIDE SEQUENCE.
RP  MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RT  "Sequence variations in the non-repetitive regions of the liver stage-
RT  specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT  isolates."
RL  Mol. Biochem. Parasitol. 71:291-294(1995).
DR  EMBL, L40888; AAC41598.1; -; Genomic_DNA.
FT  NON TER
SQ  SEQUENCE 280 AA; 32944 MW; 467085E23BABBC27 CRC64;

Query Match      35.4%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.3e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  210 RKAATKKKLERKKEHGDILAEADLYGRLEIPALTELPSENERGYIIPHOSLPODNRGNSRD 269
    |||||
DB  1 RKAATKKKLERKKEHGDVLAEDLYGRLEIPALTELPSENERGYIIPHOSLPODNRGNSRD 60

OY  270 SKEISIIKTNRESITTNVEGRDRIHKHGLEKKDGSIKPEQKEDKS 316
    |||||
DB  61 SKEISIIKTNRESITTNVEGRDRIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 11
O25851 PLAF A
ID  O25851 PLAF A PRELIMINARY; PRT; 280 AA.
AC  O25851;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Liver stage-specific antigen 1 (Fragment).
GN  Name=LSA-1;
OS  Plasmodium falciparum.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5833;
RN  NUCLEOTIDE SEQUENCE.
RP  MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA  Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RT  "Sequence variations in the non-repetitive regions of the liver stage-
RT  specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT  isolates."
RL  Mol. Biochem. Parasitol. 71:291-294(1995).
DR  EMBL, L40893; AAC41603.1; -; Genomic_DNA.
FT  NON TER
SQ  SEQUENCE 280 AA; 32927 MW; 487065EFDPCF40ADD3 CRC64;

Query Match      35.4%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.3e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  210 RKAATKKKLERKKEHGDILAEADLYGRLEIPALTELPSENERGYIIPHOSLPODNRGNSRD 269
    |||||
DB  1 RKAATKKKLERKKEHGDVLAEDLYGRLEIPALTELPSENERGYIIPHOSLPODNRGNSRD 60

OY  270 SKEISIIKTNRESITTNVEGRDRIHKHGLEKKDGSIKPEQKEDKS 316
    |||||
DB  61 SKEISIIKTNRESITTNVEGRDRIHKHGLEKKDGSIKPEQKEDKS 107

RESULT 12
O25848 PLAF A
ID  O25848 PLAF A PRELIMINARY; PRT; 280 AA.
AC  O25848;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)

```

```
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP Nucleotide sequence.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates."
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40889; AAC4159.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32939 MW; E77080385FAAD61 CRC64;

Query Match 35.1%; Score 550; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 1.8e-14;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 210 RKADTKNLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 269
DB 1 RKADTKNLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 60

QY 270 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEQKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEQKEDKS 107

RESULT 13
ID Q25853_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25853;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP Nucleotide sequence.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates."
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40834; AAB59233.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32927 MW; E6648F85FBA245 CRC64;

Query Match 35.2%; Score 550; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 1.8e-14;
Matches 105; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 210 RKADTKNLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 269
DB 1 RKADTKNLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 60

QY 270 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEQKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEQKEDKS 107

RESULT 14
ID Q25845_PLAFA PRELIMINARY; PRT; 280 AA.
```

```
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP Nucleotide sequence.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates."
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40835; AAB59234.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32938 MW; 46751C45F4DCBD33 CRC64;

Query Match 35.1%; Score 549; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 1.9e-14;
Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 210 RKADTKNLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 269
DB 1 RKADTKNLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 60

QY 270 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEQKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEQKEDKS 107

RESULT 15
ID Q25853_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25853;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP Nucleotide sequence.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Pova M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field
RT isolates."
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL, L40886; AAB59232.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32908 MW; FDEB929F34DCAT99 CRC64;

Query Match 35.1%; Score 549; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 1.9e-14;
Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 210 RKADTKNLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 269
DB 1 RKADTKNLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSLPQDNRGNSRD 60

QY 270 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEQKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEQKEDKS 107
```

Mon Dec 12 08:09:37 2005

us-09-837-344-31.rup

Page 6

Search completed: December 7, 2005, 07:20:44
Job time : 99.9815 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 06:57:58 ; Search time 102.171 Seconds
(without alignments)
1358.936 Million cell updates/sec

Title: US-09-837-344-31
Perfect score: 1563
Sequence: 1 SDLEQERRAKEXKQEQSD.....GHLEKKDSIKPEQKEDKS 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq 21:*

1:	geneseqp1980s:*
2:	geneseqp1990s:*
3:	geneseqp2000s:*
4:	geneseqp2001s:*
5:	geneseqp2002s:*
6:	geneseqp2003as:*
7:	geneseqp2003bs:*
8:	geneseqp2004s:*
9:	geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	99.1	316	2	AAR26941 P.falcipa
2	1549	99.1	493	2	AAR26944 P.falcipa
3	942.5	60.3	462	2	AAR05766 Portion o
4	771.5	49.4	318	2	AAR26943 P.falcipa
5	739	47.3	456	8	ADO21942 LSA-NRC(H
6	739	47.3	457	8	ADO21920 LSA-NRC(H
7	603	38.6	117	2	AAR26937 P.falcipa
8	331	21.2	207	3	AAB44672 Gene 33 h
9	329.5	21.1	1439	8	ADR12189 Trichomya
10	329.5	21.1	1439	8	ADU77050 Murine ha
11	329.5	21.1	1439	9	ADM00404 Epithelia
12	327.5	21.0	1162	3	AAV96255 Kaposi's
13	327.5	21.0	1162	3	AAV96255 Kaposi's
14	327.5	21.0	1162	4	AAV96255 Kaposi's
15	327.5	21.0	1162	5	ABH05621 Kaposi's
16	327.5	21.0	1162	8	ADJ65096 HHV8 late
17	327.5	21.0	1162	9	ADJ65096 HHV8 late
18	326	20.9	611	2	AAV29039 T. gondii
19	326	20.9	611	4	AAU25510 T. gondii
20	326	20.9	611	7	ADG17131 T. gondii
21	321	20.5	360	2	AAW03627 Human fol
22	321	20.5	412	2	AAW03626 Human fol
23	321	20.5	554	4	ABH59454 Drosophi
24	317.5	20.3	1898	2	AAV30795 A human t

25	317.5	20.3	1898	7	ADD48869 Human Pro
26	315.5	20.2	1584	8	ABM80133 Tumour-as
27	312	20.0	345	4	ABU53267 Human tes
28	311	19.9	621	7	ADM05724 Human pro
29	309	19.8	839	8	ADU02517 Human hum
30	306	19.6	2274	4	ABH58657 Drosophi
31	302.5	19.6	562	2	AAR70491 Leucocyto
32	301	19.3	1180	7	ADP06151 Bacterial
33	296.5	19.0	2980	8	ADP29884 Human sec
34	289	18.5	407	8	ADP49328 S pyrogen
35	289	18.5	923	4	ABG08600 Novel hum
36	288	18.4	650	7	ADB65482 Human pro
37	288	18.4	949	4	ABG15508 Novel hum
38	285	18.2	1940	8	ADP29883 Human sec
39	284	18.2	67	1	AAV81355 Sequence
40	284	18.2	1715	4	ABH58089 Drosophi
41	282	18.0	466	7	ADM05659 Human pro
42	281	18.0	658	4	ABH65632 Drosophi
43	281	18.0	800	4	ABH71459 Drosophi
44	280.5	17.9	345	8	ABO60621 Human gen
45	278	17.8	280	7	ADB64909 Human pro

ALIGNMENTS

RESULT 1					
AAR26941 standard; protein; 316 AA.					
ID	AAR26941				
XX					
AC	AAR26941:				
XX					
DT	25-MAR-2003 (revised)				
DT	08-FEB-1993 (first entry)				
XX					
DE	P.falciparum LSA-R-NR protein.				
XX					
KW	Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;				
KW	paludism; liver stage-specific antigen.				
XX					
OS	Plasmodium falciparum.				
XX					
FT	Key	Location/Qualifiers			
FT	Region	1..209			
FT		/label= repeat_region			
FT		/note= "contains 12 x 17mer repeats"			
FT	Region	210..316			
XX		/label= non-repeat_region			
PN	W09213884-A1.				
XX					
PD	20-AUG-1992.				
XX					
PF	05-FEB-1992; 92WC-FR000104.				
XX					
PR	05-FEB-1991; 91FR-00001286.				
XX					
PA	(INSP) INST PASTEUR.				
XX					
PI	Guerinmarchand C, Drulhe P;				
XX					
DR	WPI, 1992-299985/36.				
DR	N-PSDB; AAQ28115.				
XX					
PT	Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for				
PT	vaccination against, treatment of and diagnosis of malaria.				
XX					
PS	Disclosure; Fig 1; 81pp; French.				
XX					
CC	A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda				
CC	gt11 was used to transform E.coli. The expression library was screened				
CC	with human antisera against antigens of all stages of P. falciparum. The				
CC	library was rescreened with antibodies affinity- purified on a clone				

CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic USA
CC epitope. The clone with the largest insert (950 bases) encoded USA-R-NR
CC containing a 12-repeat region followed by a non-repeat region. Preferred
CC antigenic polypeptides of the invention are derived from the amino acid
CC sequence of USA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 316 AA;

Query Match 99.1%; Score 1549; DB 2; Length 316;
Best Local Similarity 99.4%; Pred. No. 1.3e-109;
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOSDLEQERRA 60
DB 1 SDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEERRAKKELQEQOSDLEQERRA 60
QY 61 KEKLEQOSDLEQERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOS 120
DB 61 KEKLEQOSDLEQERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOS 120
QY 121 DLEQERRAKKELQEQOSDLEQERRAKKELQEQOSDLEQERRAKKELQEQOSDLEQERRA 180
DB 121 DLEQERRAKKELQEQOSDLEQERRAKKELQEQOSDLEQERRAKKELQEQOSDLEQERRA 180
QY 181 EKLQEQOSDLEQERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOS 240
DB 181 EKLQEQOSDLEQERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOS 240
QY 241 IELPSENERGYIIPHOSSLPODNRGNSRDSKEISIIETKNRESITTNVGRDDIKHGHE 300
DB 241 IELPSENERGYIIPHOSSLPODNRGNSRDSKEISIIETKNRESITTNVGRDDIKHGHE 300
QY 301 EKKDGSIRPEQKEDKS 316
DB 301 EKKDGSIRPEQKEDKS 316

RESULT 2

AAR26944
ID AAR26944 standard; protein; 493 AA.

XX AAR26944;

DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)

DE P.falciparum USA gene C-terminal region.

KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KW paludism; liver stage-specific antigen.

OS Plasmodium falciparum.

Key Location/Qualifiers

FT Region 13..213
FT /label= repetitive_region
FT 214..493
FT /label= non-repetitive_region

PN WO9213884-A1.

PD 20-AUG-1992.

PF 05-FEB-1992; 92MO-FR000104.

PR 05-FEB-1991; 91FR-00001286.

PA (INSP) INST PASTEUR.

PI Guerinmarchand C, Drulhe P;

DR WPI, 1992-299985/36.

DR N-PSDB; AAQ28119.

XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.

XX Claim 2; Fig 8-10; 81pp; French.

CC The 3' part of the P.falciparum liver-stage specific antigen (USA) gene
CC codes for a polypeptide sequence which carries a T cell epitope
CC characteristic of a protein produced in hepatocytes infected with
CC P.falciparum. The polypeptide can be used in the preparation of vaccines
CC against malaria. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 493 AA;

Query Match 99.1%; Score 1549; DB 2; Length 493;
Best Local Similarity 99.4%; Pred. No. 2e-109;
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOSDLEQERRA 60
DB 5 SDLEERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOSDLEQERRA 64
QY 61 KEKLEQOSDLEQERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOS 120
DB 65 KEKLEQOSDLEQERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOS 124
QY 121 DLEQERRAKKELQEQOSDLEQERRAKKELQEQOSDLEQERRAKKELQEQOSDLEQERRA 180
DB 125 DLEQERRAKKELQEQOSDLEQERRAKKELQEQOSDLEQERRAKKELQEQOSDLEQERRA 184
QY 181 EKLQEQOSDLEQERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOS 240
DB 185 EKLQEQOSDLEQERRAKKELQEQOSDLEODRLAKEKLQEQOSDLEQERRAKKELQEQOS 244
QY 241 IELPSENERGYIIPHOSSLPODNRGNSRDSKEISIIETKNRESITTNVGRDDIKHGHE 300
DB 245 IELPSENERGYIIPHOSSLPODNRGNSRDSKEISIIETKNRESITTNVGRDDIKHGHE 304
QY 301 EKKDGSIRPEQKEDKS 316
DB 305 EKKDGSIRPEQKEDKS 320

RESULT 3

AAR05766
ID AAR05766 standard; protein; 462 AA.

XX AAR05766;

DT 25-MAR-2003 (revised)
DT 05-NOV-1990 (first entry)

DE Portion of peptide antigen to malarial sporozoite.

KW Malaria; sporozoite; vaccine; exoerythrocytic parasites; tetanus toxoid.

OS Plasmodium falciparum.

PN WO9006130-A.

PD 14-JUN-1990.

PF 30-NOV-1988; 88US-00278234.

PR 30-NOV-1988; 88US-00278234.

PA (BIOM-) BIOMEDICAL RES INST.

PI Hollingdal MR;

DR WPI, 1990-209624/27.

DR N-PSDB; AAQ05140.
 XX Novel malarial sporozoite antigenic protein - useful as vaccine against
 PT sporozoite(s) and exoerythrocytic parasites.
 XX
 PS Disclosure; Page ?; -pp; English.
 CC Antigen, preferably linked at the C-terminal to a carrier such as tetanus
 CC toxoid, may be used as a vaccine against the malarial sporozoite. (Updated
 CC on 25-MAR-2003 to correct PI field.)
 XX
 SO Sequence 462 AA;
 Query Match 60.3%; Score 942.5; DB 2; Length 462;
 Best Local Similarity 69.1%; Pred. No. 2e-63;
 Matches 210; Conservative 14; Mismatches 33; Indels 47; Gaps 3;
 QY 1 SDLEERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEERRAKKLEQEQSDLEERRA 60
 DB 180 SDLEERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEERRAKKLEQEQSDLEQRLA 239
 QY 61 KEKLEQEQSDLEERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEERRAKKLEQEQSD 103
 DB 240 KEKLEQEQSDLEERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEERRAKKLEQEQSD 299
 QY 104 DLEERRAKKLEQEQSDLEERRAKKLEQEQSDLEERRAKKLEQEQSDLEERRAK 163
 DB 300 DLEERRAKKLEQEQSDLEERRAKKLEQEQSDLEERRAKKLEQEQSDLEERRAK 359
 QY 164 EKLEQEQSDLEERRAKKLEQEQSDLEERRAKKLEQEQSDLEERRAKKLEQEQSDLE 217
 DB 360 EKLEQEQSDLEERRAKKLEQEQSDLEERRAKKLEQEQSDLEERRAKKLEQEQSDLE 419
 QY 218 LERKKEHGDILAEEDLYRLIPALPEPSENERGYIIPHOSLPQDNRGNSRDSKEISII 277
 DB 420 LERTKASTETLQE-----RQSDLEERRAKKLEQEQSDLE 455
 QY 278 KTNR 281
 DB 456 QERR 459
 RESULT 4
 ID AAR26943 standard; protein; 318 AA.
 AC AAR26943;
 XX
 AC 25-MAR-2003 (revised)
 DT 08-FEB-1993 (first entry)
 XX
 DE P. falciparum LSA N-terminal portion.
 XX
 KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
 KW paludism; liver stage-specific antigen; ss.
 OS Plasmodium falciparum.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 1..153
 FT /label= unique_5'_region
 FT 154..318
 FT /label= repetitive_region
 XX
 PN WO9213884-A1.
 XX
 PD 20-AUG-1992.
 XX
 PF 05-FEB-1992; 92WO-FR000104.
 XX
 PR 05-FEB-1991; 91PR-00001286.
 XX
 PA (INSP) INST PASTEUR.

XX
 PI Guerinmarchand C, Druihne P;
 XX WPI; 1992-299865/36.
 DR N-PSDB; AAQ28117.
 XX
 PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
 XX vaccination against, treatment of and diagnosis of malaria.
 PS Claim 11; Fig 7; 81pp; French.
 XX
 CC The 5' part of the P.falciparum liver-stage specific antigen (LSA) gene
 CC codes for a polypeptide sequence which carries a T cell epitope
 CC characteristic of a protein produced in hepatocytes infected with
 CC P.falciparum. The polypeptide can be used in the preparation of vaccines
 CC against malaria. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SO Sequence 318 AA;
 Query Match 49.4%; Score 771.5; DB 2; Length 318;
 Best Local Similarity 86.2%; Pred. No. 1.3e-50;
 Matches 163; Conservative 12; Mismatches 13; Indels 1; Gaps 1;
 QY 2 DLEERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEERRAKKLEQEQSDLEERRAK 61
 DB 131 DDKKKIKQDENRQEDLE-EKAKKLEQEQSDLEERRAKKLEQEQSDLEQRLAK 169
 QY 62 EKLEQEQSDLEERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEERRAKKLEQEQSD 121
 DB 190 EKLEQEQSDLEERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEERRAKKLEQEQSD 249
 QY 122 LEEERRAKKLEQEQSDLEERRAKKLEQEQSDLEERRAKKLEQEQSDLEERRAK 181
 DB 250 LEEERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQDRLAKKLEQEQSDLEERRAK 309
 QY 182 KLEQEQSDL 190
 DB 310 KLEQEQSDL 318
 RESULT 5
 ID ADO21942 standard; protein; 456 AA.
 AC ADO21942;
 XX
 AC 12-AUG-2004 (first entry)
 DT
 XX
 DE LSA-NRC(H) construct protein derived from Malaria parasite LSA-1.
 XX
 KW LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
 KW parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
 KW malaria parasite P. falciparum; LSA-NRC(H) construct.
 XX
 OS Plasmodium falciparum.
 OS Synthetic.
 OS
 XX
 PN WO2004044167-A2.
 XX
 PD 27-MAY-2004.
 XX
 PF 12-NOV-2003; 2003WO-US036011.
 XX
 PR 12-NOV-2002; 2002US-0425719P.
 XX
 PA (REED-) REED ARMY INST RES WALTER.
 XX
 PI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
 PI Barbosa A;
 XX
 DR WPI; 2004-420309/39.
 DR N-PSDB; ADO21941.
 XX

```

PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
PS Claim 4; SEQ ID NO 26; 90pp; English.
XX
CC The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
XX
SQ Sequence 456 AA;
Query Match 47.3%; Score 739; DB 8; Length 456;
Best Local Similarity 55.1%; Pred. No. 5,6e-48;
Matches 163; Conservative 32; Mismatches 71; Indels 30; Gaps 5;
QY 21 EODRLAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKEX 80
DB 6 EKDEIKSNLNGSSN-----SNRINIEKHKKHVSINSHSYETKNNENK 52
QY 81 LQEQSDLEQERRAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKELQEQSDLE 140
DB 53 FPDKOKELTMSNV-KVNSQTNFKSLRLNGVSENIFLKENKLNKKGKLIENHINDDDD-- 109
QY 141 QERRAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKEL 200
DB 110 KKKYIKGQDENRQEDLE-----EKAAEQSDLEQERRAKKEL-----QERRAKKEL 155
QY 201 QEQGRDLEQRKADTKKNERKKEHGDILAEDLYGRLEIPALEPSENERGYIIPHOSSLP 260
DB 156 QEQGRDLEQRKADTKKNERKKEHGDVLAEDLYGRLEIPALEPSENERGYIIPHOSSLP 215
QY 261 QDNRGNRSDSKESIIIEKTNRESITTNVGRDIIKHGLEEKDGSIKPEQKEDKS 316
DB 216 QDNRGNRSDSKESIIIEKTNRESITTNVGRDIIKHGLEEKDGSIKPEQKEDKS 271
RESULT 6
ADO21920 standard; protein; 457 AA.
XX
AC ADO21920;
XX
DT 12-AUG-2004 (first entry)
XX
DE LSA-NRC(H)Mut construct protein derived from Malaria parasite LSA-1.
XX
KM LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KM parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KM malaria parasite P. falciparum; LSA-NRC(H)Mut construct; mutant; mtein.
XX
OS Plasmodium falciparum.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 433
FT /note="Residue inserted relative to wild-type P.
FT falciparum T5 epitope"
XX
PN WO2004044167-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US036011.
XX
PR 12-NOV-2002; 2002US-0425719P.

```

```

XX
PA (REED-) REED ARMY INST RES WALTER.
XX
FI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barbosa A;
XX
DR WPI; 2004-420309/39.
DR N-PSDB; ADO21919.
XX
CC The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC within the T5 P. falciparum epitope and a C-terminal His6 tag.
XX
SQ Sequence 457 AA;
Query Match 47.3%; Score 739; DB 8; Length 457;
Best Local Similarity 55.1%; Pred. No. 5,6e-48;
Matches 163; Conservative 32; Mismatches 71; Indels 30; Gaps 5;
QY 21 EODRLAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKEX 80
DB 6 EKDEIKSNLNGSSN-----SNRINIEKHKKHVSINSHSYETKNNENK 52
QY 81 LQEQSDLEQERRAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKELQEQSDLE 140
DB 53 FPDKOKELTMSNV-KVNSQTNFKSLRLNGVSENIFLKENKLNKKGKLIENHINDDDD-- 109
QY 141 QERRAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKELQEQSDLEQERRAKKEL 200
DB 110 KKKYIKGQDENRQEDLE-----EKAAEQSDLEQERRAKKEL-----QERRAKKEL 155
QY 201 QEQGRDLEQRKADTKKNERKKEHGDILAEDLYGRLEIPALEPSENERGYIIPHOSSLP 260
DB 156 QEQGRDLEQRKADTKKNERKKEHGDVLAEDLYGRLEIPALEPSENERGYIIPHOSSLP 215
QY 261 QDNRGNRSDSKESIIIEKTNRESITTNVGRDIIKHGLEEKDGSIKPEQKEDKS 316
DB 216 QDNRGNRSDSKESIIIEKTNRESITTNVGRDIIKHGLEEKDGSIKPEQKEDKS 271
RESULT 7
AAR26937 standard; peptide; 117 AA.
XX
AC AAR26937;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P.falciparum liver-stage specific antigen non-repeat region.
XX
KM Malaria; LSA-NR; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
KM paludism.
XX
OS Plasmodium falciparum.
XX
PN WO9213884-A1.
XX
PD 20-AUG-1992.

```

```
XX 05-FEB-1992; 92MO-FR000104.
XX 05-FEB-1991; 91FR-00001286.
XX (INSP ) INST PASTEUR.
XX Guerimarchand C, Drulhe P,
XX WPI; 1992-299985/36.
XX
XX Polypeptide(6) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 4; Page 55; 81pp; French.
XX
XX A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
XX CC gII1 was used to transform E.coli. The expression library was screened
XX CC with human antisera against antigens of all stages of P. falciparum. The
XX CC library was rescreened with antibodies affinity-purified on a clone
XX CC which was able to recognise antibodies specific to the hepatic phase.
XX CC About 40 clones were detected which produced a characteristic LSA
XX CC epitope. The clone with the largest insert encoded LSA-R-NR containing a
XX CC 12-repeat region (LSA-R) followed by a non-repeat region (LSA-NR). The
XX CC sequence given here is a preferred polypeptide of the invention carrying
XX CC a T cell epitope typical of liver-stage P.falciparum. See AAR26919
XX CC R26935; AAR26937 and AAR26939. (Updated on 25-MAR-2003 to correct PN
XX CC field.)
XX
XX Sequence 117 AA;
XX
XX Query Match 38.6%; Score 603; DB 2; Length 117;
XX Best Local Similarity 100.0%; Pred. No. 2.7e-38;
XX Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 200 LQEOQRDLQORAKDPKKNLKKKKKKGGDILAEGLYGLRILPALELPSENERGYIIPHOSSL 259
XX 1 LQEOQRDLQORAKDPKKNLKKKKKKGGDILAEGLYGLRILPALELPSENERGYIIPHOSSL 60
XX
XX 260 PDNRNGNSRDSKEISIIKTNRESITTVNVEGRDIIHGKLEBKDGSIKPEOKEDKS 316
XX 61 PDNRNGNSRDSKEISIIKTNRESITTVNVEGRDIIHGKLEBKDGSIKPEOKEDKS 117
XX
XX RESULT 8
XX AAB44672
XX ID AAB44672 standard; protein; 207 AA.
XX
XX AAB44672;
XX
XX 12-FEB-2001 (first entry)
XX
XX Gene 33 human secreted protein homologous amino acid sequence #137.
XX
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX KM neotrophic; neuroprotective; antiviral; antileptergic; hepatotropic;
XX KM antidiabetic; antiinflammatory; anticancer; vulnery; anticonvulsant;
XX KM antibacterial; antifungal; antiparasitic; cardiac; gene therapy;
XX KM food additive; preservative; chromosome identification; cancer;
XX KM immune disorder; cardiovascular disorder; neurological disease;
XX KM wound healing; infectious disease.
XX
XX Homo sapiens.
XX
XX MO200058339-A2.
XX
XX 05-OCT-2000.
XX
XX 22-MAR-2000; 2000MO-US007440.
XX
XX 26-MAR-1999; 99US-0126503P.
XX 17-DEC-1999; 99US-0172409P.
XX
```

```
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsouli G;
XX WPI; 2000-594637/56.
XX
XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
XX PT the prevention, treatment and diagnosis of cancer, immune disorders,
XX PT cardiovascular disorders and neurological diseases.
XX
XX Disclosure; Page 399-400; 410pp; English.
XX
XX The polynucleotide sequences given in AAC79681 to AAC79730 encode the
XX CC human secreted proteins given in AAB44596 to AAB44645. AAB44646 to
XX CC AAB44693 represent human secreted polypeptide sequences and proteins
XX CC homologous to them, which are given in the exemplification of the present
XX CC invention. Human secreted proteins have activities based on the tissues
XX CC and cells the genes are expressed in. Examples of activities include:
XX CC cytostatic; immunosuppressive; neotrophic; neuroprotective; antiviral;
XX CC antileptergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX CC vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic; and
XX CC cardiac. The polynucleotides and polypeptides are useful for preventing,
XX CC treating or ameliorating a medical condition in e.g. humans, mice,
XX CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
XX CC can also be used as a food additive or preservative to increase or
XX CC decrease storage capabilities. The polynucleotides are useful for
XX CC chromosome identification. They are also useful as probes for diagnosing
XX CC a disorder related to the female reproductive system, particularly breast
XX CC and/or ovarian cancer. They are also useful in the gene therapy of breast
XX CC and ovarian cancer. The nucleic acids, protein, antibodies, agonists and
XX CC antagonists from the present invention are useful in the diagnosis,
XX CC treatment and prevention of: cancer; immune disorders; cardiovascular
XX CC disorders; wound healing; neurological diseases; and infectious diseases.
XX CC AAC79672 to AAC79680 and AAB44595 represent sequences used in the
XX CC exemplification of the present invention
XX
XX Sequence 207 AA;
XX
XX Query Match 21.2%; Score 331; DB 3; Length 207;
XX Best Local Similarity 31.7%; Pred. No. 2.2e-17;
XX Matches 70; Conservative 90; Mismatches 33; Indels 28; Gaps 11;
XX
XX 4 EOBERRAKEKLEQOQSDLEQDRLAKKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKE- 62
XX 6 EQQQDEBQDDEQQD-BQQQDEBQDEBQDEBQDEBQDEBQDEBQDEBQDEBQDEBQDE 64
XX
XX 63 KLEQOQSDLEQERRAKEKLEQOQSDLEQDRLAKKLEQOQSDLEQERRAKEKLEQOQSD 122
XX 65 ELEBQDELEBQ--EQLEBQDELEBQ--EQLEBQDELEBQ--EQLEBQDELEBQ--EQLEBQDE 115
XX
XX 123 EOBERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKE 182
XX 116 EEQ--EQLEBQDELEBQ--EQLEBQDELEBQ--EQLEBQDELEBQ--EQLEBQDELEBQ--EQ 163
XX
XX 183 LQEOQRDLQORAKKLEQOQSDLEQDRLAKKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKE 223
XX 164 LQEOQRDLQORAKKLEQOQSDLEQDRLAKKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKE 199
XX
XX RESULT 9
XX ADR12189
XX ID ADR12189 standard; protein; 1439 AA.
XX
XX ADR12189;
XX
XX 21-OCT-2004 (first entry)
XX
XX Trichohyalin related protein, SEQ ID 1.
XX
XX Hair growth; murine; trichohyalin.
XX
XX Mus musculus.
XX
```

PM WO2004063750-A1.
XX 29-JUL-2004.
XX
XX 07-JAN-2004; 2004WO-JP000030.
XX
XX 08-JAN-2003; 2003JP-00001891.
XX
XX (SUME) SUMITOMO ELECTRIC IND LTD.
XX
XX Takebe K, Hirai Y;
XX
XX WPI: 2004-571498/55.
XX N-PSDB; ADR12190.
XX
XX Evaluating hair-promoting activity, by using antibody that recognizes
XX proteins (antigen) of epithelial newborn follicles.
XX
XX Claim 1; SEQ ID NO 1; 39pp; Japanese.
XX
XX The present invention relates to a method (M1) for evaluating hair growth
XX -promoting activity, by using an antibody that recognizes an antigen
XX (ADR12189) of epithelial newborn follicles. Also claimed is a kit for
XX evaluating hair growth-promoting activity, comprising the antibody of
XX (M1). The present sequence is a murine trichohyalin-related protein,
XX which was used as the antigen in the method of the invention.
XX
XX Sequence 1439 AA;
SQ
Query Match 21.1%; Score 329.5; DB 8; Length 1439;
Best Local Similarity 29.8%; Pred. No. 2.3e-16;
Matches 108; Conservative 87; Mismatches 103; Indels 65; Gaps 16;
QY 2 DLQERRRAKEXX-----LQEQSDLEODRLAK--EKLQEQ-----SDLE 38
DB 602 ELQERRRRQQDRQILREEQFQREHQBARRRDETFOEEBQLQGESRRRQOEREGKFL 661
QY 39 QERRAKEXLQEQSDLEQERRAKEXX--LQEQSDLEQERRAKEXLQEQSDLEODRLAK 95
DB 662 EERQLTEREERQRRQOERERFQEEBHLQEREKELQECDRKSRQOERROOEEBQLR 721
QY 96 EKLQEQSDLEQERR--RAKEXLQEQSDLEQERR--RAKEXLQEQSDLEQERR--RAK 146
DB 722 QE-RDQFRFRREQRHLEEREEQLRDRPFRREQERRLEREEBQLRDRPFRREQERR 780
QY 147 EKLQEQSDLEQERR--RAKEXLQEQSDLEQERR--RAKEXLQEQSDLEQERRAKEXL 200
DB 781 EQLRDRPFRREQERRHLEEREEBQLRDRPFRREQERRLEREEBQLRDRPFRREQERR 839
QY 201 QEQORDLEQKADTKNLERKKEHGDILAEGLYGLRLEPAELPSENERGYIIPHOSSLP 260
DB 840 EQLRDRPFRR--EQERLEREEBQ--LDRSFRRQ--ELRRDRKTHEEERREEL 891
QY 261 QNRGNSRDSKISITTEKTNR-----SITTVNGRRDIKHGLKEKDGSI 309
DB 892 EEQGQERDRLRVEBQLRGQREEQRRQECDRKLRLEVLQLEERLRDRK--LRR 948
QY 310 EQK 312
DB 949 EQE 951
RESULT 10
ADU77050
ID ADU77050 standard; protein; 1439 AA.
XX
XX AC ADU77050;
XX
XX 10-FEB-2005 (first entry)
XX
XX Murine hair growth stimulator related protein.
XX
XX Hair growth stimulation; endocrine-gen.; cardiovascular-gen.;

KW antidiabetic; hepatotropic; antiinflammatory; gastrointestinal-gen.;
KW muscular-gen.; stimulator of hair growth; cardiovascular disease;
KW diabetes; liver disease; periodontal disease; gastrointestinal disease;
KW muscular dystrophy; cholecystitis.
OS Mus musculus.
PN WO2004101610-A1.
XX
XX 25-NOV-2004.
XX
XX 14-MAY-2004; 2004WO-JP006909.
XX
XX 16-MAY-2003; 2003JP-00138567.
XX PR 17-NOV-2003; 2003JP-00386185.
XX
XX (SUME) SUMITOMO ELECTRIC IND LTD.
XX
XX Hirai Y, Nakajima K;
XX
XX WPI: 2004-833974/82.
XX
XX Novel oligopeptide or its modified form useful for enhancing hair growth
XX or for treating cardiovascular diseases, diabetes, liver diseases,
XX periodontal disease, intestinal disease or muscular dystrophy.
XX Disclosure; SEQ ID NO 11; 80pp; Japanese.
XX
XX The invention relates to an oligopeptide or its modified form, used for
XX promoting hair growth, comprising 5-100 amino acid residues. The
XX invention also relates to a composition comprising the oligopeptide, its
XX modified form or its salt, and a hair growth stimulant comprising the
XX oligopeptide, its modified form or its salt. The oligopeptide is useful
XX for enhancing hair growth or for treating cardiovascular diseases,
XX diabetes, liver diseases, periodontal disease, intestinal disease,
XX muscular dystrophy or cholecystitis. This sequence represents a murine
XX hair growth stimulation related protein of the invention.
XX
XX Sequence 1439 AA;
SQ
Query Match 21.1%; Score 329.5; DB 8; Length 1439;
Best Local Similarity 29.8%; Pred. No. 2.3e-16;
Matches 108; Conservative 87; Mismatches 103; Indels 65; Gaps 16;
QY 2 DLQERRRAKEXX-----LQEQSDLEODRLAK--EKLQEQ-----SDLE 38
DB 602 ELQERRRRQQDRQILREEQFQREHQBARRRDETFOEEBQLQGESRRRQOEREGKFL 661
QY 39 QERRAKEXLQEQSDLEQERRAKEXX--LQEQSDLEQERRAKEXLQEQSDLEODRLAK 95
DB 662 EERQLTEREERQRRQOERERFQEEBHLQEREKELQECDRKSRQOERROOEEBQLR 721
QY 96 EKLQEQSDLEQERR--RAKEXLQEQSDLEQERR--RAKEXLQEQSDLEQERR--RAK 146
DB 722 QE-RDQFRFRREQRHLEEREEQLRDRPFRREQERRLEREEBQLRDRPFRREQERR 780
QY 147 EKLQEQSDLEQERR--RAKEXLQEQSDLEQERR--RAKEXLQEQSDLEQERRAKEXL 200
DB 781 EQLRDRPFRREQERRHLEEREEBQLRDRPFRREQERRLEREEBQLRDRPFRREQERR 839
QY 201 QEQORDLEQKADTKNLERKKEHGDILAEGLYGLRLEPAELPSENERGYIIPHOSSLP 260
DB 840 EQLRDRPFRR--EQERLEREEBQ--LDRSFRRQ--ELRRDRKTHEEERREEL 891
QY 261 QNRGNSRDSKISITTEKTNR-----SITTVNGRRDIKHGLKEKDGSI 309
DB 892 EEQGQERDRLRVEBQLRGQREEQRRQECDRKLRLEVLQLEERLRDRK--LRR 948
QY 310 EQK 312
DB 949 EQE 951

DR MPI: 2000-387829/33.
 DR N-PSDB; AAA30290.
 FT Treating or preventing a disease associated with rhodino virus infection
 PT in a mammal which includes Kaposi's Sarcoma and Primary Effusion
 PT Lymphoma.
 XX
 PS Disclosure: Fig 7; 70pp; English.
 XX
 CC The present sequence is the Kaposi's sarcoma-associated herpesvirus,
 CC (KSHV) latency-associated nuclear antigen (LANA). KSHV is also known as
 CC Human Herpes Virus 8 (HHV8) and belongs to the rhodino virus, or gamma-2
 CC herpes virus class. The LANA protein is necessary for the efficient
 CC persistence of rhadino virus DNA in mammalian cells. Persistent rhadino
 CC virus infection is implicated in a variety of diseases e.g. Kaposi's
 CC Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric
 CC Castleman's disease. In addition, KS is a common malignancy in HIV
 CC patients. KSHV persists in host cells in a latent form. One of the few
 CC genes expressed from the latent viral DNA is LANA. LANA associates with
 CC both human chromosomes and with the rhadino virus cis-acting element
 CC (RVCAE), thereby providing a tethering function: the KSHV DNA episome is
 CC "tied" to the host chromosomes. This allows the viral DNA to persist in
 CC the host cell. The present sequence may be used to screen and identify
 CC molecules that inhibit LANA interaction with RVCAE, thereby interfering
 CC with the latency cycle of this virus. Potential antiviral treatments for
 CC the above mentioned diseases may therefore be based on LANA deregulation.
 CC (Updated on 12-SEP-2003 to standardise OS field)
 CC
 XX
 SQ Sequence 1162 AA;
 Query Match 21.0%; Score 327.5; DB 3; Length 1162;
 Best Local Similarity 25.5%; Pred. No. 2,6e-16;
 Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;
 QY 4 EGERAKEXLQEQSDLEODRLAKELQ--EQQSDLEGERAKEXLQEQSDLEGERAK 61
 DB 654 EQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQ 713
 QY 62 EKLQEQQSDLEQ--ERRAKEXLQEQSDLEODRLAKEXLQEQQSDLEGERAKEXLQEQQ 119
 DB 714 QQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQ 773
 QY 120 SDLE--QERRAKE--KLQEQQSDLEQERR--AKEKLQEQQSDLEQERR--AKEKLQ 167
 DB 774 QELQEQQSDLEQEQQSDLEQEQQSDLEQEQQSDLEQEQQSDLEQEQQSDLEQEQ 833
 QY 168 EQQSDL-----EQERRAKEXLQEQ--QSDLEQ--ERRAKEXLQEQQSDLEQERRAK 219
 DB 834 EQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQ 893
 QY 220 RKKEHGDILAEIDLYGRLLEIPALELPSENERGYIIPHQSSLPQD--NRGNSRDSKEISII 276
 DB 894 EVEEQEQLLEVEEQEQLLEVEEQEQLLEVEEQEQLLEVEEQEQLLEVEEQEQLLE 953
 QY 277 EKTNRRESITTNVEG 290
 DB 954 VVSTHEQIASSPPG 967
 RESULT 13
 AAAY58500
 ID AAAY58500 standard; protein; 1162 AA.
 XX
 AC AAAY58500;
 XX
 DT 06-AUG-2003 (revised)
 DT 10-APR-2000 (first entry)
 XX
 DE HHV8 ORF 73 protein, SEQ ID NO:21.
 XX
 KM HHV8; detection; diagnosis; Kaposi's sarcoma; AIDS; immunogen; antigen.
 XX
 OS Human herpesvirus 8.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 96
 FT /label= unknown
 XX
 EN WO9961909-A2.
 XX
 PD 02-DEC-1999.
 XX
 PE 26-MAY-1999; 99WO-US011407.
 XX
 FR 26-MAY-1998; 98US-0086695P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Pau C;
 DR MPI: 2000-097142/08.
 XX
 PS New methods and compositions for the detection of human herpesvirus.
 XX
 PS Claim 2; Page 59-62; 68pp; English.
 XX
 CC Sequences AAAY58480-Y58532 represent immunogenic polypeptides derived from
 CC human herpes virus type 8 (HHV8, a gammaherpesvirus). HHV8 plays an
 CC important role in the pathogenesis of AIDS-related Kaposi's sarcoma. The
 CC invention relates to a novel method of detecting the presence of human
 CC herpesvirus 8 in a biological sample using peptides representative of
 CC dominant antigenic regions of HHV8. The method comprises contacting one
 CC or more isolated, immunogenic HHV8 peptides with an antibody-containing
 CC biological sample, and detecting the formation of a complex between the
 CC peptide and the antibody. The presence of a peptide-antibody complex
 CC indicates the presence of human herpesvirus 8. The detection of HHV8
 CC infection can be used to diagnose AIDS-associated Kaposi's sarcoma. The
 CC HHV8-specific antibodies are useful therapeutically when for the passive
 CC immunisation of a human against HHV8 infection, thereby reducing HHV8
 CC related disease. The detection assays are highly specific, sensitive and
 CC accurate. Early detection and treatment of Kaposi's sarcoma could
 CC diminish the severity of symptoms related to AIDS and the sensitive
 CC techniques could reduce erroneous characterisations of skin disorders.
 CC Previous assays for HHV8 antibodies such as immunofluorescence assays,
 CC immunoblots and enzyme immunoassays lack the sensitivity and accuracy
 CC needed for reliable diagnosis of Kaposi's sarcoma. Further advantages of
 CC the assays are that reproducible results are obtained and the method is
 CC suitable for rapid throughput and screening of samples economically.
 CC (Updated on 06-AUG-2003 to correct OS field.)
 CC
 XX
 SQ Sequence 1162 AA;
 Query Match 21.0%; Score 327.5; DB 3; Length 1162;
 Best Local Similarity 25.5%; Pred. No. 2,6e-16;
 Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;
 QY 4 EGERAKEXLQEQSDLEODRLAKELQ--EQQSDLEGERAKEXLQEQQSDLEGERAK 61
 DB 654 EQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQ 713
 QY 62 EKLQEQQSDLEQ--ERRAKEXLQEQSDLEODRLAKEXLQEQQSDLEGERAKEXLQEQQ 119
 DB 714 QQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQ 773
 QY 120 SDLE--QERRAKE--KLQEQQSDLEQERR--AKEKLQEQQSDLEQERR--AKEKLQ 167
 DB 774 QELQEQQSDLEQEQQSDLEQEQQSDLEQEQQSDLEQEQQSDLEQEQQSDLEQEQ 833
 QY 168 EQQSDL-----EQERRAKEXLQEQ--QSDLEQ--ERRAKEXLQEQQSDLEQERRAK 219
 DB 834 EQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQ 893
 QY 220 RKKEHGDILAEIDLYGRLLEIPALELPSENERGYIIPHQSSLPQD--NRGNSRDSKEISII 276
 DB 894 EVEEQEQLLEVEEQEQLLEVEEQEQLLEVEEQEQLLEVEEQEQLLEVEEQEQLLE 953

OY 277 EKTRESITTNVEG 290
DB 954 VVSTHEQIASSPPG 967

RESULT 14
AAB62331
ID AAB62331 standard; protein; 1162 AA.
XX
AC AAB62331;
XX
DT 06-AUG-2003 (revised)
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of KSHV tethering protein LANA.
XX
KM Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KM Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KW KSHV; latency-associated nuclear antigen; LANA.
XX
OS Human herpesvirus 8.
XX
PN WO200125484-A2.
XX
PD 12-APR-2001.
XX
PF 29-SEP-2000; 2000WO-US026908.
XX
PR 01-OCT-1999; 99US-00410399.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Robertson ES, Cotter MA;
XX
DR WPI; 2001-281736/29.
DR N-PSDB; AAF82901.
XX
PT A composition for use in gene therapy comprises an expression vector that
PT includes a nucleic acid sequence encoding a nucleic acid binding protein.
XX
PS Disclosure; Fig 9B; 60pp; English.
XX
CC The invention provides a composition comprising nucleic acid, histone H1
CC protein and expression vector operably encoding a protein suitable
CC for tethering the nucleic acid to the histone H1 protein, where the
CC tethering protein is LANA. The composition is useful in aiding the
CC retention of the viral DNA in the host cell. The viral vector encodes a
CC protein suitable for tethering DNA to Histone H1. Methods for screening
CC for compounds which are agonistic or antagonistic for the tethering of
CC viral proteins to histone H1 and DNA binding sites are useful for
CC developing the method of viral transfer. The composition has applications
CC to gene therapy, including the treatment of multiple sclerosis,
CC Parkinson's disease, Huntington disease and diabetes. The present
CC sequence represents the amino acid sequence of the Kaposi's sarcoma
CC associated herpesvirus (human herpesvirus 8) latency-associated nuclear
CC antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 1162 AA;

Query Match 21.0%; Score 327.5; DB 4; Length 1162;
Best Local Similarity 25.5%; Pred. No. 2.6e-16;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

OY 4 EGBRRAKELQEQSDLEQDRLAKEKIQ--EQQSDLEGBRRAKELQEQSDLEGBRRAK 61
DB 654 EQQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQ 713
OY 62 EKLQEQSQDLEQ--ERRAKELQEQSQDLEQDRLAKEKIQEQSQDLEGBRRAKELQEQ 119
DB 714 QQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQ 773
OY 120 SDELE-GBRRAKE---KLEQEQSQDLEGBR---AKKELQEQSQDLEGBR---AKKELQ 167

DB 774 QLELDQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLE 833
OY 168 EQQSDLE-----EQBRRAKELQEQ--QSDLEQ--ERRAKELQEQQDRLEQKADTKKNLE 219
DB 834 EQQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQ 893
OY 220 RKKEHGDILAEADLYGRLEIPALIEPSENEGYIYIPHQSSLPQD---NRGNSRDSKEISII 276
DB 894 EVEEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEBFEQEQLEB 953
OY 277 EKTRESITTNVEG 290
DB 954 VVSTHEQIASSPPG 967

RESULT 15
ABB05621
ID ABB05621 standard; protein; 1162 AA.
XX
AC ABB05621;
XX
DT 25-APR-2002 (first entry)
XX
DE Kaposi's sarcoma-associated herpesvirus LANA protein.
XX
KM Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAE; PEL;
KM KSHV terminal repeat; rhadino virus cis acting element; episome;
KM primary effusion lymphoma; latency-associated nuclear antigen;
KM gene therapy; gene transfer.
XX
OS Human herpesvirus 8.
XX
PN US6322792-B1.
XX
PD 27-NOV-2001.
XX
PF 21-APR-1999; 99US-00298568.
XX
PR 19-NOV-1998; 98US-0109422P.
XX
PA (KIEF/) KIEF E D.
XX
PI Kieff ED, Ballesgas ME, Kaye KW;
XX
DR WPI; 2002-153769/20.
DR N-PSDB; ABA93487.
XX
PT System for episomal retention of plasmids in mammalian cells, useful in
PT gene therapy, comprises rhadinoviral LANA and RVCAE sequences.
XX
PS Disclosure; Fig 7; 27pp; English.
XX
CC The present invention describes a system (A) for maintaining a plasmid as
CC an episome in mammalian cells, comprising the rhadinoviral sequence LANA
CC (latency-associated nuclear antigen) of 3409 base pairs (see ABA93487,
CC S1) expressed in the cell, and the rhadinoviral sequence RVCAE
CC (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, S2)
CC present in the plasmid. Also described is a method for maintaining a
CC closed circular DNA in a cell by expressing (S1) in the cells and having
CC (S2) as a cis-acting and maintenance sequence in the DNA. (A) is
CC particularly used in gene therapy (or other gene transfer applications)
CC that uses mammalian cells in which LANA is expressed. (A) improves
CC persistence of gene therapy vectors in cells. The present sequence
CC represents Kaposi's sarcoma-associated herpesvirus (KSHV, also called
CC human herpesvirus 8) LANA protein, which is used in the exemplification
CC of the present invention
XX
SQ Sequence 1162 AA;

Query Match 21.0%; Score 327.5; DB 5; Length 1162;
Best Local Similarity 25.5%; Pred. No. 2.6e-16;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:16:04 ; Search time 31.3811 Seconds
(without alignments)
832.525 Million cell updates/sec

Title: US-09-837-344-31

Sequence: 1 SDLEGERRAKEKEEQGQSDL.....GHLEKKDSIRKEQKEDKS 316

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/H COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/PTCUS COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfill1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	316	2	US-08-098-327E-31 Sequence 31, Appl
2	1563	100.0	316	2	US-08-462-625-31 Sequence 31, Appl
3	603	38.6	117	2	US-08-098-327E-20 Sequence 20, Appl
4	603	38.6	117	2	US-08-462-625-20 Sequence 20, Appl
5	554	35.4	107	2	US-08-098-327E-19 Sequence 19, Appl
6	554	35.4	107	2	US-08-462-625-19 Sequence 19, Appl
7	344	22.0	608	2	US-09-270-767-32937 Sequence 32937, A
8	344	22.0	608	2	US-09-270-767-48154 Sequence 48154, A
9	327.5	21.0	1162	1	US-08-728-323A-2 Sequence 2, Appli
10	327.5	21.0	1162	2	US-09-298-568-2 Sequence 2, Appli
11	327.5	21.0	1162	2	US-09-410-359-2 Sequence 2, Appli
12	327.5	21.0	1162	2	US-09-894-273-2 Sequence 2, Appli
13	326	20.9	611	2	US-09-216-393B-81 Sequence 81, Appl
14	317.5	20.3	1898	1	US-08-056-200-94 Sequence 94, Appl
15	317.5	20.3	1898	1	US-08-800-644-94 Sequence 94, Appl
16	317.5	20.3	1898	1	US-09-538-092-1280 Sequence 1280, Ap
17	301	19.3	1180	2	US-09-543-681A-6436 Sequence 6436, Ap
18	288	18.4	650	2	US-10-104-047-3636 Sequence 3636, Ap
19	278	17.8	280	2	US-10-104-047-3063 Sequence 3063, Ap
20	277	17.7	735	2	US-10-164-595-80 Sequence 80, Appl
21	277	17.7	784	2	US-10-164-595-79 Sequence 79, Appl
22	277	17.7	843	2	US-10-164-595-54 Sequence 54, Appl
23	275	17.6	779	2	US-10-164-595-56 Sequence 56, Appl
24	272	17.4	1564	2	US-10-144-198-2 Sequence 4, Appli
25	272	17.4	1564	2	US-10-144-198-4 Sequence 4, Appli
26	270.5	17.3	919	2	US-09-949-016-6954 Sequence 6954, Ap
27	268	17.1	798	2	US-10-104-047-3335 Sequence 3335, Ap

28	265	17.0	676	2	US-09-270-767-44900	Sequence 44900, A
29	260	16.6	1402	2	US-09-248-796A-14503	Sequence 14503, A
30	259	16.6	613	2	US-09-543-681A-6012	Sequence 6012, Ap
31	255.5	16.3	1960	2	US-09-538-092-1077	Sequence 1077, Ap
32	255.5	16.3	1960	2	US-09-549-016-10872	Sequence 10872, Ap
33	255	16.3	683	6	5210183-3	Patent No. 5210183
34	253	16.2	2101	1	US-08-466-390-4	Sequence 4, Appli
35	253	16.2	2101	1	US-08-470-950-4	Sequence 4, Appli
36	253	16.2	2101	1	US-08-467-781-4	Sequence 1, Appli
37	253	16.2	2101	1	US-08-195-487-4	Sequence 4, Appli
38	253	16.2	2101	1	US-08-483-924-4	Sequence 4, Appli
39	253	16.2	2101	2	US-09-452-294-1	Sequence 32, Appli
40	253	16.2	2101	2	US-09-296-662-32	Sequence 4, Appli
41	253	16.2	2101	4	PCT-US93-06160-4	Sequence 1078, Ap
42	250.5	16.0	1976	2	US-09-538-092-1078	Sequence 7646, Ap
43	250	16.0	2107	2	US-09-949-016-7646	Sequence 7647, Ap
44	250	16.0	2107	2	US-09-949-016-7647	Sequence 33, Appli
45	250	16.0	2115	2	US-09-296-662-33	

ALIGNMENTS

RESULT 1
US-08-098-327E-31
Sequence 31, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SRO ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-31
Query Match 100.0%; Score 1563; DB 2; Length 316;

Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
Db 1 SDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
QY 61 KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQS 120
Db 61 KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQS 120
QY 121 DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 180
Db 121 DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 180
QY 181 EKLQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 240
Db 181 EKLQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 240
QY 241 IELPENERGYIIPHQSSLPQDNRGNSRDSKEISIEKTNRSEITTNVEGRDIIKHGLE 300
Db 241 IELPENERGYIIPHQSSLPQDNRGNSRDSKEISIEKTNRSEITTNVEGRDIIKHGLE 300
QY 301 EKKGSIRPEQEKDS 316
Db 301 EKKGSIRPEQEKDS 316

RESULT 2

US-08-462-625-31
Sequence 31, Application US/08462625
Patent No. 6319502

GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. PALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-31

Query Match 100.0%; Score 1563; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
Db 1 SDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRA 60
QY 61 KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQS 120
Db 61 KEKLEQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQS 120
QY 121 DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 180
Db 121 DLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 180
QY 181 EKLQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 240
Db 181 EKLQEQSDLEQERRAKKLEQEQSDLEQDRLAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAK 240
QY 241 IELPENERGYIIPHQSSLPQDNRGNSRDSKEISIEKTNRSEITTNVEGRDIIKHGLE 300
Db 241 IELPENERGYIIPHQSSLPQDNRGNSRDSKEISIEKTNRSEITTNVEGRDIIKHGLE 300
QY 301 EKKGSIRPEQEKDS 316
Db 301 EKKGSIRPEQEKDS 316

RESULT 3

US-08-098-327E-20
Sequence 20, Application US/08098327E
Patent No. 6270771

GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. PALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-20

Query Match 38.6%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.9e-42;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 LQEQQRLEQRKADTKNLERKKEHGDILAEGLYGRLEIPALIEPSENERGYIIPHOSSL 259
DB 1 LQEQQRLEQRKADTKNLERKKEHGDILAEGLYGRLEIPALIEPSENERGYIIPHOSSL 60

QY 260 PDNRGNSRDSKEISIEKTNRESITTNVEGRDIIHGHLSEKKDGSIKPEQKEDKS 316
DB 61 PDNRGNSRDSKEISIEKTNRESITTNVEGRDIIHGHLSEKKDGSIKPEQKEDKS 117

RESULT 4
US-08-462-625-20
Sequence 20, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
NUMBER OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
CLASSIFICATION: 424
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-20

Query Match 38.6%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4.9e-42;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 LQEQQRLEQRKADTKNLERKKEHGDILAEGLYGRLEIPALIEPSENERGYIIPHOSSL 259
DB 1 LQEQQRLEQRKADTKNLERKKEHGDILAEGLYGRLEIPALIEPSENERGYIIPHOSSL 60

QY 260 PDNRGNSRDSKEISIEKTNRESITTNVEGRDIIHGHLSEKKDGSIKPEQKEDKS 316
DB 61 PDNRGNSRDSKEISIEKTNRESITTNVEGRDIIHGHLSEKKDGSIKPEQKEDKS 117

RESULT 5
US-08-098-327E-19
Sequence 19, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
NUMBER OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-19

Query Match 35.4%; Score 554; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 4.2e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 RKADTKNLERKKEHGDILAEGLYGRLEIPALIEPSENERGYIIPHOSSL PDNRGNSRD 269
DB 1 RKADTKNLERKKEHGDILAEGLYGRLEIPALIEPSENERGYIIPHOSSL PDNRGNSRD 60

QY 270 SKEISIIKTNRESITTNVEGRDIIKHGHEKKDGIKPEOKEDKS 316
Db 61 SKEISIIKTNRESITTNVEGRDIIKHGHEKKDGIKPEOKEDKS 107

RESULT 6

US-08-462-625-19
Sequence 19, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Machis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-19

Query Match 35.4%; Score 554; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 4, 2e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 210 RKADTKKNIERKKEHGDIIAEDLYGRLEIPALPSPENRGYYIPIQSSIPQDNNGNSRD 269
Db 1 RKADTKKNIERKKEHGDIIAEDLYGRLEIPALPSPENRGYYIPIQSSIPQDNNGNSRD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIKHGHEKKDGIKPEOKEDKS 316
Db 61 SKEISIIKTNRESITTNVEGRDIIKHGHEKKDGIKPEOKEDKS 107

RESULT 7

US-09-270-767-32937

Sequence 32937, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32937
LENGTH: 608
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-32937

Query Match 22.0%; Score 344; DB 2; Length 608;
Best Local Similarity 31.7%; Pred. No. 3, 5e-20;
Matches 71; Conservative 77; Mismatches 68; Indels 8; Gaps 4;

QY 4 BOERRAKEXLOEQSDLEQDRLAKEXLOEQSDLEQERRAKEXLOEQSDLEQERRAKEX 63
Db 381 BEERRKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEER 440
QY 64 LOEQSDLEQERRAKEXLOEQSDLEQDRLAKEXLOEQSDLEQERRAKEXLOEQSDLE 123
Db 441 REBARK-EEERRKAEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKE 496
QY 124 QERRA-KEKLOEQSDLEQERRAKEXLOEQSDLEQERRAKEXLOEQSDLEQERRAKEX 182
Db 497 EERRKDEERRERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEER 556
QY 183 LOEQSDLEQERRAKEXLOEQSDLEQ--RKADTKKNIERKE 223
Db 557 RDEERRERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKE 600

RESULT 8
US-09-270-767-48154
Sequence 48154, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 48154
LENGTH: 608
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-48154

Query Match 22.0%; Score 344; DB 2; Length 608;
Best Local Similarity 31.7%; Pred. No. 3, 5e-20;
Matches 71; Conservative 77; Mismatches 68; Indels 8; Gaps 4;
QY 4 BOERRAKEXLOEQSDLEQDRLAKEXLOEQSDLEQERRAKEXLOEQSDLEQERRAKEX 63
Db 381 BEERRKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEER 440
QY 64 LOEQSDLEQERRAKEXLOEQSDLEQDRLAKEXLOEQSDLEQERRAKEXLOEQSDLE 123
Db 441 REBARK-EEERRKAEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKE 496
QY 124 QERRA-KEKLOEQSDLEQERRAKEXLOEQSDLEQERRAKEXLOEQSDLEQERRAKEX 182
Db 497 EERRKDEERRERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEERKEER 556
QY 183 LOEQSDLEQERRAKEXLOEQSDLEQ--RKADTKKNIERKE 223

Db 557 ROEERRRERREERREERREERREERREERREERREERREERREERRE 600

```
RESULT 9
US-08-728-323A-2
; Sequence 2, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Ruess, James J.
; APPLICANT: Edelman, Jaidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,323A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MSC/SKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-728-323A-2
```

```
Query Match 21.0%; Score 327.5; DB 1; Length 1162;
Best Local Similarity 25.5%; Pred. No. 1.6e-18;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

QY 4 EOERRAKEKLOEQSDLEQDRLAKKLO--EQSDLEQERRAKEKLOEQSDLEQERRAK 61
Db EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDE 713
QY 62 EKLQEQSDLEQ--ERRAKEKLOEQSDLEQDRLAKKLOEQSDLEQERRAKEKLOEQ 119
Db 714 QQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDE 773
QY 120 SDLE-QERRAKE--KLOEQSDLEQERR--AKKLOEQSDLEQERR--AKKLOEQ 167
Db 774 QELDDEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLE 833
QY 168 EQQSDL-----EOERRAKEKLOEQ--QSDLEQ--ERRAKEKLOEQORDLEQKADTKKYLE 219
Db 834 EQQLEEQLEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVE 893
QY 220 RKKEHGDILADLYGRLEIPALIEPSENERGYIIPHOSSLPOD--NRGNSRDSKEISII 276
Db 894 EEEQEQLEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVE 953
QY 277 EKTRESITTNVEG 290
```

Db 954 VVSTHEQIASSPG 967

```
RESULT 10
US-09-298-568-2
; Sequence 2, Application US/09298568
; Patent No. 6322792
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Balleslas, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/298,568
; CURRENT FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,422
; EARLIER FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-2
```

```
Query Match 21.0%; Score 327.5; DB 2; Length 1162;
Best Local Similarity 25.5%; Pred. No. 1.6e-18;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

QY 4 EOERRAKEKLOEQSDLEQDRLAKKLO--EQSDLEQERRAKEKLOEQSDLEQERRAK 61
Db EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDE 713
QY 62 EKLQEQSDLEQ--ERRAKEKLOEQSDLEQDRLAKKLOEQSDLEQERRAKEKLOEQ 119
Db 714 QQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDE 773
QY 120 SDLE-QERRAKE--KLOEQSDLEQERR--AKKLOEQSDLEQERR--AKKLOEQ 167
Db 774 QELDDEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLEEQLE 833
QY 168 EQQSDL-----EOERRAKEKLOEQ--QSDLEQ--ERRAKEKLOEQORDLEQKADTKKYLE 219
Db 834 EQQLEEQLEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVE 893
QY 220 RKKEHGDILADLYGRLEIPALIEPSENERGYIIPHOSSLPOD--NRGNSRDSKEISII 276
Db 894 EEEQEQLEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVEEQVE 953
QY 277 EKTRESITTNVEG 290
Db 954 VVSTHEQIASSPG 967

RESULT 11
US-09-410-399-2
; Sequence 2, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Cotter, Murray A.
; TITLE OF INVENTION: Method to Inhibit or Enhance the Binding of Viral DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT
```

```

; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-389-2
Query Match
Best Local Similarity 21.0%; Score 327.5; DB 2; Length 1162;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

QY 4 EERRAKKLEQEQSDLEQDRLAKEKLEQ--EQQSDLEERRAKKLEQEQSDLEQERRAK 61
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 654 EQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQDE 713
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 EKLQEQSDLEQ--ERRAKKLEQEQSDLEQDRLAKEKLEQEQSDLEQERRAKKLEQEQ 119
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 714 QQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQDE 773
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 120 SDLE-QERRAKE--KLEQEQSDLEQERR--AKKLEQEQSDLEQERR--AKKLEQ 167
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 774 QELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELE 833
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 168 EQQSDLE-----EQERRAKKLEQEQ--QSDLEQ--ERRAKKLEQEQORDLEQKADTKKLE 219
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 834 EQEQELEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQELE 893
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 220 RKKEGDIADLYGRLEIPALPELPSENERGYIIPHOSSLPOD--NRGNSRDSKEISII 276
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 894 EYEEQEQELEQEQEQELEQEQELEQEQEQEQEQEQEQEQEQEQEQEQEQELE 953
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 277 EKTNRRESITTNVEG 290
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 954 VVSTHEQIASSPPG 967

RESULT 12
US-09-894-273-2
; Sequence 2, Application US/09894273
; Patent No. 6756203
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballesetas, Mary E.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/894,273
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/109,422
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-2
Query Match
Best Local Similarity 21.0%; Score 327.5; DB 2; Length 1162;
Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10;

QY 4 EERRAKKLEQEQSDLEQDRLAKEKLEQ--EQQSDLEERRAKKLEQEQSDLEQERRAK 61
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 654 EQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQDE 713
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 EKLQEQSDLEQ--ERRAKKLEQEQSDLEQDRLAKEKLEQEQSDLEQERRAKKLEQEQ 119
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 714 QQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQDE 773
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 120 SDLE-QERRAKE--KLEQEQSDLEQERR--AKKLEQEQSDLEQERR--AKKLEQ 167
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 774 QELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELEQEQELE 833
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 168 EQQSDLE-----EQERRAKKLEQEQ--QSDLEQ--ERRAKKLEQEQORDLEQKADTKKLE 219
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```

DB 834 EQEQELEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQELE 893
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 220 RKKEGDIADLYGRLEIPALPELPSENERGYIIPHOSSLPOD--NRGNSRDSKEISII 276
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 894 EYEEQEQELEQEQEQELEQEQELEQEQEQEQEQEQEQEQEQEQEQEQEQELE 953
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 277 EKTNRRESITTNVEG 290
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 954 VVSTHEQIASSPPG 967

RESULT 13
US-09-216-393B-81
; Sequence 81, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milnausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDI PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 81
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-216-393B-81
Query Match
Best Local Similarity 20.9%; Score 326; DB 2; Length 611;
Matches 89; Conservative 99; Mismatches 99; Indels 54; Gaps 12;

QY 4 EQERR--AKKLEQEQSDLEQDRLAKEKLEQEQSDLEQERR--AKKLEQEQSDLEQEQ 57
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 292 EERRRAREEKEKERRRQEQEERERRRVEEKEKERRQEQEERERRRVEEKKARQREDERER 351
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 58 RRAKKEKLEQEQSDLEQERR--AKKLEQEQSDLEQDRLAKEKLEQEQSDLEQERRAKE 113
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 352 RRVEEKKARQEQEERERRRVEEKKARQEQEERERRRVEEKKARQEQEERERRRVE 411
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 114 KLEQEQSDLEQERRAK--EKLQEQSDLEQEQ--RRAKKEKLEQEQSDLEQERRAKEKLEQ 168
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 412 EKKARQREEEERERRRVEEKKARQREEEERERRRVEEKKARQREEEERERRRVEEKE 471
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 169 ---QQSDLEQERRAKKLEQEQSDLEQERR--AKKLEQEQORDLEQKADTKKLEBK 221
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 472 RERQERERRRRRVEEKEKERRQEQEERERRRVEEKEKERRQEQEERERRRVEEKKARQRO 531
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 222 KEHGDILADLYGRLEIPALPELPSENERGYIIPHOSSLPODNRGNSRDSK-EISII EKTN 280
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 532 EEEG-----RERQNG--EEEREREREQEQERELKTRIVELQREH 569
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 281 RESITTN--NVEGRARDI-----KGHLEKKKGSIKPEOK 312
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 570 AESVETWKEQGERERHLLTOPMERKLLHAFEESSRTVLQER 610
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 14
US-08-056-200-94
; Sequence 94, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichinella and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Knobbe, Martens, Olson & Bear
;; STREET: 620 Newport Center Drive, Sixteenth Floor
;; CITY: Newport Beach
;; STATE: CA
;; COUNTRY: U.S.A.
;; ZIP: 92660
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/056,200
;; FILING DATE: 30-APR-1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fedrick, Michael F.
;; REGISTRATION NUMBER: 36,799
;; REFERENCE/DOCKET NUMBER: NIH054,001A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (714) 760-0404
;; TELEFAX: (714) 760-9502
;; INFORMATION FOR SEQ ID NO: 94:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1898 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-056-200-94

Query Match 20.3%; Score 317.5; DB 1; Length 1898;
Best Local Similarity 29.7%; Pred. No. 1.9e-17;
Matches 105; Conservative 77; Mismatches 116; Indels 55; Gaps 14;
QY 4 EERRAKKLEQEQSDLEODRLAKKLEQEQSDLEQERRAKKLEQEQSDL-----E 55
DB 458 EERR-EGRLKREQE--RDWLKREETERH--EQERRKQQLKRDQERRERRLKLE 512
QY 56 QERRAKKLEQEQSDLEQERRAKKLEQEQSDLEODRLAKKLEQEQSDLEQERRAKKLE 115
DB 513 EERRQEQERRRQQLRREQERRRQQLKRE--EERRQQLRNSQQLRREQERRLEQL 569
QY 116 -QEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSD-L 173
DB 570 KREERKLEQERRRQQLRREQERRRQQLKREERKQQLRREQERRRQQLKREERLE 629
QY 174 EERRAKKLEQEQ-----SDLEQERRAKKLEQEQDLEQ--KADTKKLER 220
DB 630 EERRRDRRLKREPERERRRHELLKSEQERRRHEQLRREQERRRQQLKREERLEQ 689
QY 221 --KKEGDILAEADLYGLE-----IPATLPSENERGYIYHQSSL---PODN 263
DB 690 RLKREHEERRRQQLAEQERRRQQLKREERKSRIPKQWQLESD--ARQSVLLEAPQAG 746
QY 264 RGNRDSKEISIIKTNRESITTVNVEGRDILKHLKLEKDGSIKPEKEDKS 316
DB 747 RAAPQEQE---EKRRRESELQWQERRRAHQEQERRRQQLRFTWQQAEEKS 795
RESULT 15
US-08-800-644-94
; Sequence 94, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methode of Using Same
; NUMBER OF SEQUENCES: 117

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Knobbe, Martens, Olson & Bear
;; STREET: 620 Newport Center Drive, Sixteenth Floor
;; CITY: Newport Beach
;; STATE: CA
;; COUNTRY: U.S.A.
;; ZIP: 92660
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/800,644
;; FILING DATE: 14-FEB-1997
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/056,200
;; FILING DATE: 30-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fedrick, Michael F.
;; REGISTRATION NUMBER: 36,799
;; REFERENCE/DOCKET NUMBER: NIH054,001A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (714) 760-0404
;; TELEFAX: (714) 760-9502
;; INFORMATION FOR SEQ ID NO: 94:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1898 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-800-644-94

Query Match 20.3%; Score 317.5; DB 1; Length 1898;
Best Local Similarity 29.7%; Pred. No. 1.9e-17;
Matches 105; Conservative 77; Mismatches 116; Indels 55; Gaps 14;
QY 4 EERRAKKLEQEQSDLEODRLAKKLEQEQSDLEQERRAKKLEQEQSDL-----E 55
DB 458 EERR-EGRLKREQE--RDWLKREETERH--EQERRKQQLKRDQERRERRLKLE 512
QY 56 QERRAKKLEQEQSDLEQERRAKKLEQEQSDLEODRLAKKLEQEQSDLEQERRAKKLE 115
DB 513 EERRQEQERRRQQLRREQERRRQQLKRE--EERRQQLRNSQQLRREQERRLEQL 569
QY 116 -QEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSDLEQERRAKKLEQEQSD-L 173
DB 570 KREERKLEQERRRQQLRREQERRRQQLKREERKQQLRREQERRRQQLKREERLE 629
QY 174 EERRAKKLEQEQ-----SDLEQERRAKKLEQEQDLEQ--KADTKKLER 220
DB 630 EERRRDRRLKREPERERRRHELLKSEQERRRHEQLRREQERRRQQLKREERLEQ 689
QY 221 --KKEGDILAEADLYGLE-----IPATLPSENERGYIYHQSSL---PODN 263
DB 690 RLKREHEERRRQQLAEQERRRQQLKREERKSRIPKQWQLESD--ARQSVLLEAPQAG 746
QY 264 RGNRDSKEISIIKTNRESITTVNVEGRDILKHLKLEKDGSIKPEKEDKS 316
DB 747 RAAPQEQE---EKRRRESELQWQERRRAHQEQERRRQQLRFTWQQAEEKS 795
Search completed: December 7, 2005, 07:22:12
Job time : 32.3811 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:18:34 ; Search time 82.4665 Seconds
(without alignments)
1601.063 Million cell updates/sec

Title: US-09-837-344-31

Sequence: 1 SDLEGRRAKEKLEQEQSD.....GHLEKKDSIKPEQKEDKS 316

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:*

- 1: /cgn2_6/ptcdat/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptcdat/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptcdat/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptcdat/1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptcdat/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptcdat/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	316	3	US-09-837-344-31 Sequence 31, Appl
2	1563	100.0	316	3	US-09-900-963-31 Sequence 31, Appl
3	1494	95.6	1909	5	US-10-732-923-3341 Sequence 3341, Ap
4	603	38.6	117	3	US-09-837-344-20 Sequence 20, Appl
5	603	38.6	117	3	US-09-900-963-20 Sequence 20, Appl
6	554	35.4	107	3	US-09-837-344-19 Sequence 19, Appl
7	554	35.4	107	3	US-09-900-963-19 Sequence 19, Appl
8	355	22.7	1003	5	US-10-732-923-3319 Sequence 3319, Ap
9	329.5	21.1	1439	5	US-10-754-079-157 Sequence 157, App
10	328	21.0	1407	5	US-10-732-923-3359 Sequence 3359, Ap
11	327.5	21.0	1162	3	US-09-894-273-2 Sequence 2, Appl1
12	327.5	21.0	1162	4	US-10-294-804-2 Sequence 2, Appl1
13	327.5	21.0	1162	4	US-10-194-046-2 Sequence 2, Appl1
14	326	20.9	611	3	US-09-321-193-81 Sequence 81, Appl
15	326	20.9	611	4	US-10-321-856-81 Sequence 81, Appl
16	321	20.5	554	6	US-11-097-143-5154 Sequence 5154, Ap
17	320	20.5	350	4	US-10-101-487-58 Sequence 58, Appl
18	320	20.5	350	5	US-10-939-988-58 Sequence 58, Appl
19	311	19.9	621	4	US-10-108-260A-4409 Sequence 4409, Ap
20	310.5	19.9	1549	4	US-10-732-923-3358 Sequence 3358, Ap
21	307	19.6	240	4	US-10-101-487-75 Sequence 75, Appl
22	307	19.6	240	5	US-10-939-988-75 Sequence 75, Appl
23	306	19.6	2274	6	US-11-097-143-2763 Sequence 2763, Ap
24	289	18.5	923	5	US-10-450-763-38959 Sequence 38959, A
25	288	18.4	650	4	US-10-104-047-3636 Sequence 3636, Ap
26	288	18.4	949	5	US-10-450-763-45867 Sequence 45867, A
27	284	18.2	1715	6	US-11-097-143-1059 Sequence 1059, Ap

28	282	18.0	466	4	US-10-108-260A-4344 Sequence 4344, Ap
29	281	18.0	658	6	US-11-097-143-23688 Sequence 23688, A
30	281	18.0	800	6	US-11-097-143-41169 Sequence 41169, A
31	280.5	17.9	345	4	US-10-029-386-34255 Sequence 34255, A
32	278	17.8	280	4	US-10-104-047-3063 Sequence 3063, Ap
33	277	17.7	735	4	US-10-144-194A-109 Sequence 109, App
34	277	17.7	735	5	US-10-491-566-109 Sequence 109, App
35	277	17.7	735	5	US-10-717-665-80 Sequence 80, App
36	277	17.7	784	4	US-10-144-194A-108 Sequence 108, App
37	277	17.7	784	5	US-10-491-566-108 Sequence 108, App
38	277	17.7	784	5	US-10-717-665-79 Sequence 79, Appl
39	277	17.7	843	4	US-10-144-194A-84 Sequence 84, Appl
40	277	17.7	843	5	US-10-491-566-84 Sequence 84, Appl
41	277	17.7	843	5	US-10-717-665-54 Sequence 54, Appl
42	275	17.6	779	4	US-10-144-194A-86 Sequence 86, Appl
43	275	17.6	779	5	US-10-491-566-86 Sequence 86, Appl
44	275	17.6	779	5	US-10-717-665-56 Sequence 56, Appl
45	273	17.5	507	4	US-10-425-114-47285 Sequence 47285, A

ALIGNMENTS

RESULT 1
US-09-837-344-31
Sequence 31, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-837-344-31

Query Match 100.0%; Score 1563; DB 3; Length 316;
Best Local Similarity 100.0%; Pred. No. 8.5e-85;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDLEERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 60
DB 1 SDLEERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 60
QY 61 KEKLEQOQSDLEQERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQ 120
DB 61 KEKLEQOQSDLEQERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQ 120
QY 121 DLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 180
DB 121 DLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 180
QY 181 EKLQEQSDLEQERRAKEKLEQOQSDLEQDRKADTKKLERKKEHGDILAEIDL YGRLEIIPA 240
DB 181 EKLQEQSDLEQERRAKEKLEQOQSDLEQDRKADTKKLERKKEHGDILAEIDL YGRLEIIPA 240
QY 241 IELPENERGYIIPHOSSLPODNRGNSRDSKEISIEKTNRRESITTNVGRRDHKGHLE 300
DB 241 IELPENERGYIIPHOSSLPODNRGNSRDSKEISIEKTNRRESITTNVGRRDHKGHLE 300
QY 301 EKKDGSIKPEQEKDGS 316
DB 301 EKKDGSIKPEQEKDGS 316

RESULT 2

US-09-900-963-31
Sequence 31, Application US/09900963
Publication No. US20030064075A1

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

HEPATITIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE

OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/900,963

FILING DATE: 10-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/098,327

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-045

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 316 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-900-963-31

Query Match 100.0%; Score 1563; DB 3; Length 316;
Best Local Similarity 100.0%; Pred. No. 8.5e-85;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDLEERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 60
DB 1 SDLEERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 60
QY 61 KEKLEQOQSDLEQERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQ 120
DB 61 KEKLEQOQSDLEQERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQ 120
QY 121 DLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 180
DB 121 DLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 180
QY 181 EKLQEQSDLEQERRAKEKLEQOQSDLEQDRKADTKKLERKKEHGDILAEIDL YGRLEIIPA 240
DB 181 EKLQEQSDLEQERRAKEKLEQOQSDLEQDRKADTKKLERKKEHGDILAEIDL YGRLEIIPA 240
QY 241 IELPENERGYIIPHOSSLPODNRGNSRDSKEISIEKTNRRESITTNVGRRDHKGHLE 300
DB 241 IELPENERGYIIPHOSSLPODNRGNSRDSKEISIEKTNRRESITTNVGRRDHKGHLE 300
QY 301 EKKDGSIKPEQEKDGS 316
DB 301 EKKDGSIKPEQEKDGS 316

RESULT 3

US-10-732-923-3341

Sequence 3341, Application US/10732923

Publication No. US20050108791A1

GENERAL INFORMATION:

APPLICANT: Edgerton, Michael D

TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

FILE REFERENCE: 38-15(52796)C

CURRENT FILING DATE: 2003-12-10

PRIOR APPLICATION NUMBER: 10/310,154

PRIOR FILING DATE: 2002-12-04

NUMBER OF SEQ ID NOS: 24149

SEQ ID NO 3341

LENGTH: 1909

TYPE: PRT

ORGANISM: Plasmodium falciparum

US-10-732-923-3341

Query Match 95.6%; Score 1494; DB 5; Length 1909;
Best Local Similarity 95.9%; Pred. No. 5.1e-80;
Matches 303; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 SDLEERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 60
DB 1421 SDLEERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 1480
QY 61 KEKLEQOQSDLEQERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQ 120
DB 1481 KEKLEQOQSDLEQERRAKEKLEQOQSDLEQDRLAKEKLEQOQSDLEQERRAKEKLEQOQ 1540
QY 121 DLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 180
DB 1541 DLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRAKEKLEQOQSDLEQERRA 1600
QY 181 EKLQEQSDLEQERRAKEKLEQOQSDLEQDRKADTKKLERKKEHGDILAEIDL YGRLEIIPA 240

Db 1601 EKQEQORDLEQRKAKKQEQORDLEQRKADTKNLERKKEHGVLAEDLYGRLEIPA 1660
QY 241 IELPSENERGYIIPHOSSLPODNRGNRSRDSKSIIEKTNRESITTVNVEGRDIIHKHGLE 300
Db 1661 IELPSENERGYIIPHOSSLPODNRGNRSRDSKSIIEKTNRESITTVNVEGRDIIHKHGLE 1720
QY 301 EKKDGSIKPEQKEDKS 316
Db 1721 EKKDGSIKPEQKEDKS 1736

RESULT 4
US-09-837-344-20
; Sequence 20, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01266
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-837-344-20

Query Match 38.6%; Score 603; DB 3; Length 117;
Best local similarity 100.0%; Pred. No. 1.3e-28; Indels 0; Gaps 0;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 LQEQORDLEQRKADTKNLERKKEHGVLAEDLYGRLEIPAIELPSENERGYIIPHOSSL 259
Db 1 LQEQORDLEQRKADTKNLERKKEHGVLAEDLYGRLEIPAIELPSENERGYIIPHOSSL 60
QY 260 PODNRGNRSRDSKSIIEKTNRESITTVNVEGRDIIHKHGLEKDGSIKPEQKEDKS 316

Db 61 PODNRGNRSRDSKSIIEKTNRESITTVNVEGRDIIHKHGLEKDGSIKPEQKEDKS 117

RESULT 5
US-09-900-963-20
; Sequence 20, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-900-963-20

Query Match 38.6%; Score 603; DB 3; Length 117;
Best local similarity 100.0%; Pred. No. 1.3e-28; Indels 0; Gaps 0;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 LQEQORDLEQRKADTKNLERKKEHGVLAEDLYGRLEIPAIELPSENERGYIIPHOSSL 259
Db 1 LQEQORDLEQRKADTKNLERKKEHGVLAEDLYGRLEIPAIELPSENERGYIIPHOSSL 60
QY 260 PODNRGNRSRDSKSIIEKTNRESITTVNVEGRDIIHKHGLEKDGSIKPEQKEDKS 316
Db 61 PODNRGNRSRDSKSIIEKTNRESITTVNVEGRDIIHKHGLEKDGSIKPEQKEDKS 117

RESULT 6
US-09-837-344-19
; Sequence 19, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine

DRUIHE, Pierre
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-837-344-19
Query Match 35.4%; Score 554; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 9.2e-26;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 210 RKADTKKNIKKERKKGHDILADLYGRLEIPALIEPSENERGYIIPHQSSLPQDNKGNRSD 269
DB 1 RKADTKKNIKKERKKGHDILADLYGRLEIPALIEPSENERGYIIPHQSSLPQDNKGNRSD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107
RESULT 7
US-09-900-963-19
Sequence 19, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIHE, Pierre
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404

CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-900-963-19
Query Match 35.4%; Score 554; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 9.2e-26;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 210 RKADTKKNIKKERKKGHDILADLYGRLEIPALIEPSENERGYIIPHQSSLPQDNKGNRSD 269
DB 1 RKADTKKNIKKERKKGHDILADLYGRLEIPALIEPSENERGYIIPHQSSLPQDNKGNRSD 60
QY 270 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 316
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107
RESULT 8
US-10-732-923-3319
Sequence 3319, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 3319
LENGTH: 1003
TYPE: PRT
ORGANISM: Kapoisi's sarcoma-associated herpesvirus
US-10-732-923-3319
Query Match 22.7%; Score 355; DB 5; Length 1003;
Best Local Similarity 25.3%; Pred. No. 3.8e-13;
Matches 79; Conservative 116; Mismatches 63; Indels 54; Gaps 14;
QY 4 EOERRAKELQEQSD--LEQDRLAKERLQEQSDLEERRAKERLQEQSDLEERRAK 61

[illegible]

```

RESULT 9
US-10-754-079-157
; Sequence 157, Application US/10754079
; Publication No. US2005008602A1
GENERAL INFORMATION:
APPLICANT : Hiral, Yohel
APPLICANT : Takebe, Kyoko
TITLE OF INVENTION: OLIGOPEPTIDES FOR PROMOTING HAIR GROWTH
FILE REFERENCE: 467392000222
CURRENT APPLICATION NUMBER: US/10/754, 079
CURRENT FILING DATE: 2004-01-07
PRIOR APPLICATION NUMBER: US 10/155, 922
PRIOR FILING DATE: 2002-05-23
PRIOR APPLICATION NUMBER: PCT/JPO1/04691
PRIOR FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: JP 2000-166903
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: JP 2001-347340
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: JP 2001-347338
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: JP 2001-371175
PRIOR FILING DATE: 2001-12-05
PRIOR APPLICATION NUMBER: JP 2001-371336
PRIOR FILING DATE: 2001-12-05
PRIOR APPLICATION NUMBER: JP 2003-1891
PRIOR FILING DATE: 2003-01-08
NUMBER OF SEQ ID NOS: 162
SOFTWARE: FaecesEQ for Windows Version 4.0
SEQ ID NO 157
LENGTH: 1439
TYPE: PRF
ORGANISM: Mus musculus
US-10-754-079-157

Query Match      21.1%; Score 329.5; DB 5; Length 1439;
Basic Local Similarity    29.8%; Pid. No. 1,7e+11;
Matches 108; Conservative 87; Mismatches 103; Indels   65; Gaps   1d

Oy       2 DLQERRAKKK-----LQQQSDDLEDDRLLK---EKLOEQQ-----SDLE 38
           ||| |::||::||::||::||:||||:
Db        602 ELBERRRRROQRDRLIEEFQPFRHNRKRARRDFPFQEEDLOGESSRRROGKERCKTLE 666
           ::||::||::||::||:||||:

Oy       39 QERBAKEKLQQSDDLQERRAKE--LQQQSDEERRRAKEYKLQSQSDDLEDRIAL 95
           ::||::||::||::||:||||:
Db        662 EEROLTREREFORRQOEERFOEEBHHQLDEREXELRJOCDSKRESEORRQOREEOURLR 722
           ::||::||::||::||:||||:

Oy       96 EKLOEQSSDLQER---RAKEYLKQSQSDDLEQER---RAKEYLKQSQSDDLEQER---RAK 146
```

```

Db      722  QE-ROGRFRREOGRHLEEREEOQLRDRPSRREOGRHQRENEEOQLDRPSRRQOEHQOREE 760
Oy      147  EKLQEOQSDLEOGR---RAKELQEOQSDLEOGR---RAKELQEOQSDLEOGRRAEKL 200
Db      781  EOARDPSRREOGRHOERREEOQLRDRPSRREOGRLEEREEOQLDRPSRRQOEKH--QER 839
Oy      201  QEOORPLEORKADTKTKLERKKENHDILAEDYGLERIPALEPSENERGYIPIHQSLP 260
Db      840  EOQLRRDRPSRR--EOERLREEREO---LRDSFRREQ---ELARDKKFHHEERRELE 891
Oy      261  QDNRGNSRDSKEISIIIEKTRE-----SITTVBGRDIIHKLHEEKDQSIKP 309
Db      892  EEOGRGDRRLRVEEOQLRGOREEOORROECDRKHLRELYOEALIEERLDRDK---LRR 948
Oy      310  EOK 312
Db      949  EOE 951

```

```

RESULT 10
US-10-732-923-3359
; Sequence 3359, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3359
; LENGTH: 1407
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-732-923-3359

Query Match      21.0%; Score 328; DB 5; Length 1407;
Best Local Similarity 29.4%; Pred. No. 2e-11;
Matches    88; Conservative   77; Mismatches   92; Indels    42; Gaps    10;

QY      4 EOERRAKEKLOE-----QQSDLEODRLAKETLQEOQSDLEOERR-----AKKLQEOQSDL 54
         |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      150 EEOKKRRFRPOHYSROYRDKEQ-RLORQELERRABEEOLRRKGRDAEFIEEQARR 208
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      55 EOBRRAKETLQEOQSDLEOERRAKEK-LQEOQSDLEODRLAKETLQEOQSDLEOERRAKE 113
         :||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      209 REQQLRKLEEBEQRREREQEHRALQEBEEOLTKORRRREEPREQO---QRRELE 264
         :||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      114 KLOEQQSDLEOERRAKEKTLQEOQSDLEOE-----BRAKETLQEOQSDLEOERRAKEKTL 166
        ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      265 EIRREERQRLBEERRREQQLRRERQR-LQEBERRREQQLRLEIEINEREGRLQEBERRREQRL 323
        ::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      167 QEOQSDLEOERRAKEKTLQEOQSDLEOERRAKEKTLQEOQSDLEOQRADTKNLERKKEGHD 226
        :::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      324 EQERRREQQLRKLEIEIREREREQRLQEBERRREQLLAEFVR-----EQAREGE 370
        :::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      227 ILAEDLYARLEIPALELPSENERGYIYP--HOSSLPODNRGNRSDEKSIIIEKTRE 282
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      371 SLTRRWOROLE---SEAGARSKVYSRREROEOSLTRDODERRORQERERELBEOARRQ 426
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 11
US-09-894-273-2
; Sequence 2, Application US/09894273
; Publication No. US20040037847A1
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballesta, Mary E.
; APPLICANT: Kaye, Kenneth W.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: RHADINO DNA TO MEDIANTE EFFICIENT EPISOME PERSISTENCE

```

```

FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/894,273
CURRENT FILING DATE: 2001-06-28
PRIORITY APPLICATION NUMBER: US 60/109,422
PRIOR FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1162
TYPE: PRT
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-2

```

Query Match	21.0%;	Score 327.5;	DB 3;	Length 1162;
Best Local Similarity	25.5%;	Pred. No. 1.8e-11;		
Matches	80;	Conservative 104;	Mismatches 103;	Indels 27;
			Gaps	10

Dy 4 E0ERAKKEKJQEOOSDLEODRLAKEKJQ--E0OSSDEOERRAKEKLQEQSSDLQEERRAK 61
 ||:: ::::| | :||::| |::| |::| |:
Db 654 EQQQQEQQQEQQQEQQQEQQQEQQQEQQQEQQQEQQQEQQQEQQQEQQQEQQQEQQQEQQQE 713

Page 98 of 100

```
Oy      62 EKLQEQQSDLEQ--ERRAKELQEQQSDEQLRLAKELQEQQSDLEPERRAKEKLEQEQ 113
        ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      714 QQDEQQQDDEQQQDDFEQQQDDFEQQQDDFEQQQDDFEQQEQQEEQQEQQEQLERQE 773
```

```
QY      120 SDLE-QERRAKE---KXQEQSDLEGERR-----AKEXQ 1674  
       :|||: : |||: |||: :|||: |||: :|||:  
DB      774 QELEQEQELEEEQEQELEEEQEQELEEEQEQELEEEQEQELE 833
```

Qy 168 EQQSD-----EQERAKEXKQEQ-QSDLEQ-ERAKEXKQEQQRDRQKADTKNLE 219

Db 834 EQEQLEEQVEEQEQVEEQEQEQEQLEEQVEEQEQEQEQEQEQLEEQVEEQEQEQLE 833

Qy 220 RKKHHDLADLYGRLEIPALELPSENERGYIIPHOSSLPD--NNGNSKUSKEISLI 2/8

Db 894 EVEBOEQLLEVEEEOQQELLEVEEEOQQGVQEQEQTVEEPIILHOSSSEDEMEVDYP 953

QY	277	ERINKESITINVEG	290
	:		:
	:	:	:
DB	954	VVSTHEQIASSPPG	967

RESULT 12
US-10-294-804-2
: Sequence 2, Application US/10294804

```

; FUGILATION NO.: 0200000007081
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; APPLICANT: Cotter, Murray A.
```

```

: TITLE OF INVENTION: to Genomic Host DNA
: FILE REFERENCE: UM-03778
: CURRENT APPLICATION NUMBER: US/10/294,804

```

; PRIOR APPLICATION NUMBER: US/09/410,399
 ; PRIOR FILING DATE: 1999-10-01
 ; NUMBER OF SEQ ID NOS: 6

```

; SEQ ID NO 2
;
; LENGTH: 1162
;
; TYPE: PRT
;

```

US-10-294-804-2	
Query Match	21.0%; Score 327.5; DB 4; Length 1162;

Matches 80; Conservative 104; Mismatches 103; Indels 27; Gaps 10

D6 654 EQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQEQQQDEQE 711
QY 62 EKLEQQSDLEQ--ERRAKEKLEQQSDELDRIAKEKLEQQSDLEQRRAKEKLEEQ 119

[illegible]

RESULT 13
US-10-194-046-2

;; Sequence 2, Application US/10194046
;; Publication No. US20040248081A1
;; GENERAL INFORMATION:
;; APPLICANT: Robertson, Erle S.

```

; TITLE OF INVENTION: LANA Viral Pict
; FILE REFERENCE: UM-07142
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2002-07-11
; CREATION AND/OR INVENTION NUMBER: 00/410 2

```

```

; PRIOR APPLICATION NUMBER: 09/410,3
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 3

```

```

; SEQ ID NO 2
;
; LENGTH: 1162
;
; TYPE: PRT
;
; ORGANISM: Kaposi's sarcoma-associ
nt: 10 104 046-3

```

Query Match	21.0%	Score
Best Local Similarity	25.5%	Percentage
Marches	80	Conservative
	104	

Matches	80; conservative 104;
QY	4 EQERRAKETQEQQSDLEQDRLL :: :::: :: ::
Dh	654 RQODEQODEEQODEEQODEEQ

Dd 62 EKLEQQSDDEQ--ERRAKEKLC
 :: : :: | : : | : : :
Db 714 QDEEQQDDEEQQDDDEEQQDDDEEQ

```

Db      774 QELEDQEQLEEFQEQLEFEQEQ
Qy      120 SDLE-QERRAKE---KLQEQQS
        :|||: :| :|||:

```

```

Qy      168 EQQSDJ-----EQERPAKEKLC
          ||: :|      |||: :|:
Db      834 EQEQLEEQVEEQEQVEEQEQE

```

Qy 220 RKKHGDILAEADLYGRLEIPAT
: | : | : :
Db 894 EVEEEOEQLEEEVEEEOEQLEEE

Qy	277	EKTRESITNVEG	290
	:		:
Db	954	VSTHEQIASSPG	967

RESULT 14
US-09-216-393-01

; Sequence 81, Application US/092163
; Patent No. US2001001447A1

[illegible]

Standing Slices

herpesvirus

```

7.7.5; DB 5; length 1162;
1.1.8e-11;
atches 103; Indels 27; Gaps 10

```

```

-----E00SDLBQERAKKLTQEQ00SDLBQERAK 61
      ||| | : : : ||| | : || :
DE0000DE0000DE0000DE0000DE0000DE00 713

```

LEEDORLAKELQEOQSLEGERAKELOEQO 115
:: :: |||| : || :: :: ||:
OOODEOQQDEQEQQEEOEQOEEOEFLEEQE 773

AKELQEQSDLEOERR---AKEKLO 167
::|::|::|::|
QELBEQOELEBEOQELEBEOQELEBEOQELE 833

```

DLEQ-ERAKELQEQORDLEQRKADTKNLE 219
||| : | : : ||| : : ||
ELEFEQEQQEQQEEQEEQELLEVEEQEELE 893

```

ENERGYI PHOSLPD--NRGNSPDSKEISII 276
:::|:::|:|:|:
QOQGEQOQEQETVEEPIILHGSSSEDEMEVDYP 953

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:21:30 ; Search time 7.29792 Seconds

(without alignments)
241.806 Million cell updates/sec

Title: US-09-837-344-31

Sequence: 1 SDLEQERRAKEKLOEQSD.....GHLEKKDGSIRPKQKDKS 316

Scoring table:

BLOSUM62

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0 %
Maximum Match 100 %

Listing first 45 summaries

Database :

Published Applications AA New:*

- 1: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US66_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241	15.4	793	7	US-11-060-914-2 Sequence 2, Appli
2	235	15.0	885	6	US-10-793-626-1660 Sequence 1660, Ap
3	234	15.0	1388	6	US-10-821-234-1143 Sequence 1143, Ap
4	226	14.5	795	6	US-10-770-726-49 Sequence 49, Appl
5	218.5	14.0	1404	6	US-10-878-556A-169 Sequence 169, Appl
6	212.5	13.6	718	6	US-10-878-556A-97 Sequence 97, Appl
7	209.5	13.4	1410	6	US-10-821-234-1050 Sequence 1050, Ap
8	193.5	12.4	1189	7	US-11-074-176-134 Sequence 134, Ap
9	188.5	12.1	531	7	US-11-060-914-4 Sequence 4, Appl1
10	186.5	11.9	1410	6	US-10-878-556A-136 Sequence 136, App
11	186.5	11.9	1586	6	US-10-821-234-901 Sequence 901, App
12	186	11.9	155	6	US-10-952-535A-17 Sequence 17, Appl
13	186	11.9	284	6	US-10-821-234-1332 Sequence 1632, Ap
14	185	11.8	664	6	US-10-878-556A-44 Sequence 44, Appl
15	184.5	11.7	558	6	US-10-467-657-694 Sequence 694, App
16	183.5	11.7	248	6	US-10-878-556A-175 Sequence 175, App
17	183	11.7	1565	6	US-10-467-657-2704 Sequence 2704, Ap
18	181	11.6	145	6	US-10-952-535A-19 Sequence 19, Appl
19	180.5	11.5	672	7	US-11-000-463-455 Sequence 455, App
20	179	11.5	121	6	US-10-952-535A-14 Sequence 14, Appl
21	178.5	11.4	1299	6	US-10-821-234-1145 Sequence 1145, Ap
22	176.5	11.3	1043	6	US-10-821-234-1055 Sequence 1055, Ap
23	175	11.2	989	6	US-10-821-234-975 Sequence 975, App
24	174	11.1	691	6	US-10-131-826A-16 Sequence 16, Appl
25	174	11.1	703	6	US-10-821-234-963 Sequence 963, App

26	173.5	11.1	898	7	US-11-099-691-7 Sequence 7, Appl1
27	170	10.9	1122	6	US-10-467-657-6112 Sequence 6112, Ap
28	168	10.7	676	7	US-11-135-855-28 Sequence 28, Appl
29	168	10.7	717	7	US-11-135-855-29 Sequence 29, Appl
30	165.5	10.6	465	6	US-10-878-556A-116 Sequence 116, App
31	160.5	10.3	495	6	US-10-821-234-1154 Sequence 1154, Ap
32	158	10.1	600	6	US-10-878-556A-43 Sequence 43, Appl
33	157.5	10.1	1616	6	US-10-821-234-1497 Sequence 1497, Ap
34	155	9.9	585	6	US-10-878-556A-42 Sequence 42, Appl
35	154.5	9.9	418	6	US-10-878-556A-78 Sequence 78, Appl
36	150.5	9.6	459	6	US-10-821-234-896 Sequence 896, App
37	149	9.5	882	7	US-11-012-762-34 Sequence 34, Appl1
38	148	9.5	1618	6	US-10-984-645-2 Sequence 2, Appl1
39	146	9.3	440	7	US-11-008-172-1059 Sequence 1059, App
40	146	9.3	3674	7	US-11-000-463-454 Sequence 454, App
41	145	9.3	1095	6	US-10-793-626-3154 Sequence 3154, Ap
42	144	9.2	1596	7	US-11-060-005-4 Sequence 4, Appl1
43	143.5	9.2	431	7	US-11-074-176-132 Sequence 132, App
44	143.5	9.2	639	6	US-10-821-234-907 Sequence 907, App
45	143	9.1	747	6	US-10-821-234-1662 Sequence 1662, Ap

ALIGNMENTS

```
RESULT 1
US-11-060-914-2
; Sequence 2, Application US/11060914
; Publication No. US20050261184A1
; GENERAL INFORMATION:
; APPLICANT: Kaufman, Paul L
; APPLICANT: Geiger, Benjamin
; APPLICANT: Berthodsky, Alexander
; APPLICANT: Borrias, Teresa
; TITLE OF INVENTION: Method for Treating Glaucoma
; FILE REFERENCE: 960296.00128
; CURRENT APPLICATION NUMBER: US/11/060.914
; CURRENT FILING DATE: 2005-02-18
; PRIOR APPLICATION NUMBER: US 60/545,722
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 60/545,723
; PRIOR FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-060-914-2

Query Match      15.4%; Score 241; DB 7; Length 793;
Best Local Similarity 25.4%; Pred. No. 6.6e-06;
Matches 85; Conservative 73; Mismatches 105; Indels 72; Gaps 12;

4 EQERRAKEKLOEQ-----SDLEQDLRAKEKLOEQSD-----LEQERRAKE 45
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 ENKKDEKKEKEEKEEKEPRGSGIGENOVWVEEKTTSOETTVMSLKNQISSEBPKE 241
KLOEQSD-----LEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQDLRAKEKLOE 100
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
242 EERQGSSEISHHKMEEDKERAEARAEARER---ERIKAQDDKIADERARIEAE 299
QOSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQER 160
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
300 KAAAOEERRRAE---ERRRMEERKEAAE---ERRRIKEBEKRAE---ERRRIKEBEK 350
RAKKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOEQSDLEQERRAKEKLOE 220
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
351 PAAE---ERQRIKEBEKRAE---ERQRRABEEBKAKVEQKQKQLEK---KHAMQE 401
KKEGDLIAELVGRLEIPALEPSENERGYYPHQSGLPDQNRGNSRDSKEISITKTN 280
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
402 TKIGEKVEQKIEBKW-----NKKAQEDLQTAVALKKQ 436
```

QY 281 RESITTVNGBRDHKGHEKKDSIRPEKEDK 315
DB 437 GEKGTQVQAKRE---KLQEDKPTPKKEIKDEK 467

RESULT 2

US-10-793-626-1660
; Sequence 1660, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMELLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1660
; LENGTH: 885
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1660

Query Match 15.0%; Score 235; DB 6; Length 885;
Best Local Similarity 25.2%; Pred. No. 1.6e-07;
Matches 89; Conservative 86; Mismatches 90; Indels 88; Gaps 17;

QY 4 EGERRAKELQEQ-----QSDLEQDRIAKELQEQSDLEQERRAKELQEQS 52
DB 501 EQGTQA-EQLSEQYFASQOYNNLKEQVHHELELRLKTQEAHLKNEH---EEFEFEKN 556
QY 53 DLEQERRAKELQEQSDLEQERRAKELQEQSDLEQ-----DLAKE-----KLQEQ 101
DB 557 DGYQSDSKSEYTLKEKNHLE---IQQQLKQLESIDERYTLKSEKASTHOTQOQLHK 613
QY 102 QSDL-----EGERRAKELQEQSDLEQERRAKELQEQSDLEQERRAKELQEQ 153
DB 614 QSDLAIVKKEIKSQKQYERLQKQSDSERQKLEVENKIKLTPSDMKDAFEKLEKQ 672
QY 154 SDLEQERRAKELQEQSDLEQERRAKELQEQSDLE-----QERRAKE- 198
DB 673 --IQQENVRONLQQLSEIKQKQKDLNEKIEINESQLQKHQDILSIENHYQDIKAKQS 730
QY 199 -----KLQEQORDLEQKADTKKLERKEHGDILADLYGRLEIPAI 241
DB 731 KLDVLINHAIDLNDTYQLTVERARMEYSDFTIDILRKVKLTKMTIBEL-GPVVLNMI 789
QY 242 ELPSR-NEGRGYVPHQSSLPQDNRGNSRDSKEI--SIIETKRESITTVNEGR 291
DB 790 EQEELNERTYPLNEQ-----RTDLREKELTLEQILHMDKE-----VEGR 830

RESULT 3

US-10-821-234-1143
; Sequence 1143, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07

; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1143
; LENGTH: 1388
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-821-234-1143

Query Match 15.0%; Score 234; DB 6; Length 1388;
Best Local Similarity 23.4%; Pred. No. 2.8e-07;
Matches 92; Conservative 86; Mismatches 98; Indels 118; Gaps 17;

QY 7 RRAKELQEQSDLEQDRILA-KEKLEQEQSDLEQERRAKELQEQ-----RAKELQEQ----- 49
DB 594 QREKELQREKELQKVKABEERLRQAKREKREKRIIQEHQIKKTVRERLEQIKKTE 653
QY 50 -----QSDLEQERRAKELQEQSDLEQERRAKELQEQ-----EKL 81
DB 654 LGAKAFKDIDIEDLELDPDFIMAKQVQLKEKKELEQLERLKNQEKKIDYFERRARLEBI 713
QY 82 QEQSDLEQDRIAKELQEQSD-----LEQERRAKELQEQSDLEQERRA 128
DB 714 PLIKAYEQRKMDLWEQEQEERITTMQLERKALEKONMSMLEDRDLFWRLKKA 773
QY 129 KEKLEQEQSDLEQERRAKELQEQSDLEQERRA---KEKLEQEQSDLEQERRAKEL 182
DB 774 RQSVYERLQEQEERLEAERNNRLEERKRQREKRIITYRKEKEEQ-----RAEQ 827
QY 183 LQEQSDLEQERRAKELQEQORDLEQKADTKK--NLERKEHGDILADLYGRLEIPA 240
DB 828 MKEREEREERARAKR--EELREYQER---VKLEVERKKRQRE-----LEIEE 873
QY 241 IELPSRNEGRYVPHQSSLPQDNRGNSRDSKISIIETKRESITTVNEGR----- 291
DB 874 RRRREER--RLGSSLSRKDSRWGRDS-----EGTWKGPADSEMRGPPREKWR 925
QY 292 ---RDHKGHL-----BEKQDSIKPE 310
DB 926 RQGRDEBDSHRDRERPRRLGDDDRPSLRPD 959

RESULT 4

US-10-770-726-49
; Sequence 49, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 795
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-770-726-49

Query Match 14.5%; Score 226; DB 6; Length 795;
Best Local Similarity 22.6%; Pred. No. 4.8e-07;
Matches 76; Conservative 78; Mismatches 105; Indels 78; Gaps 12;

QY 5 EERRAKELQEQSDLEQDRIAKELQEQSDLEQERRAKELQEQSDLEQERRAKEL 64
DB 89 QQMSRKEKVHNR-----KQKREKRRHSHSABGKARVKEK-----EHRKRGRHE 139
QY 65 QEQSDLEQERRAKELQEQSDLEQDRIAKELQEQSDLEQERRAKELQEQSDLEQ 124
DB 140 EDDKARERWERQKREMAHESRRDRRL--QLERK--RRERGRQEQEQORQOKER 194


```
Db      1092 LEEERERRERREKLGQWRQTOSLNPAPSPLTAAQWKEKSTLQRPQETVIRE----- 1146
Qy      46 KLQEQOSDLEQERRAKEKLQEQOSDLEQ-----ERRAKEKLQEQOSDLEQRLAKE 96
Db      1147 -LQPOQOPPTERRDLOYTVTSKEELSSGTVCPPDPWKKDAKEKLEKQOQMHIVMLKE 1205
Qy      97 KLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSD 156
Db      1206 -LQEQOSKPKRBAESDRLKMLQEQOKLQEQOSKQKEDDEBEEDDDVDITMLMQR-L 1263
Qy      157 EERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKE 213
Db      1264 EERRAKRTMPALISVLDLQDEBERRRQQLQEMRK-READBARQEQ-EERRAQEER--- 1318
Qy      214 TKKLERKKEHGDILADLYGRLEIPALEPSENERGYIPIH-----QSSLPQD 262
Db      1319 TKRDAEKKRRQ---EEGYSRLEAEERRRRQHDAAARRLLEPAPGLCRPPLPRD 1368
```

```
RESULT 8
US-11-074-176-134
; Sequence 134, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Rusebell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; FILE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-134
```

```
Query Match      12.4%; Score 193.5; DB 7; Length 1189;
Best Local Similarity 21.0%; Pred. No. 5.1e-05;
Matches 74; Conservative 89; Mismatches 105; Indels 85; Gaps 14;
```

```
Qy      4 EERRAKEKLQEQOSDLEQ-----EODRLAKEKLQEQOSDLEQERRAKEKL 44
Db      172 QOKEEQONOLKKTQDNLIRINDLVKELSRLEPLNEQSSLAKE-YKFOKSGDLKUKSL 230
Qy      45 -----EKLQEQOSDLEQ-----RRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQ 92
Db      231 APEIENINQORREIQKSAQDKNKLKLDDEVKDSQAATYQKRAEYKRLDER-DHTQNK 289
Qy      93 LAK--EKLQEQOSDLEQ-----EERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKE 147
Db      290 LKLSKDLSEINSLQMAEQSRQFDDATKEEYKN-QVQKQKQNLVQLKADLDLKLKEXK 347
Qy      148 KLQEQOSDLEQ-----RAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSD 191
Db      348 KLQEQOSDLEQ-----RAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSD 406
Qy      192 EERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSD 251
Db      407 ---RRQADITTYQGVDSKQLTDAQKLE-----QALIEGKGLDKRQK-- 447
Qy      252 YIPQSSLPQDNFGNSRDSKEISIT-----EKTNRESITTVNVEGRDILHKGH 298
Db      448 ---EQNAIVIRKQNNQNLTELNLKQVVAEERNELEKVARREALVNIQKRH 497
```

```
RESULT 9
US-11-060-914-4
; Sequence 4, Application US/11060914
; Publication No. US20050261184A1
; GENERAL INFORMATION:
; APPLICANT: Kaufman, Paul L
; APPLICANT: Geiger, Benjamin
; APPLICANT: Bershady, Alexander
; APPLICANT: Borras, Teresa
; TITLE OF INVENTION: Method for Treating Glaucoma
; FILE REFERENCE: 960296.00128
; CURRENT APPLICATION NUMBER: US/11/060,914
; CURRENT FILING DATE: 2005-02-18
; PRIOR APPLICATION NUMBER: US 60/545,722
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 60/545,723
; PRIOR FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-060-914-4
```

```
Query Match      12.1%; Score 188.5; DB 7; Length 531;
Best Local Similarity 22.1%; Pred. No. 4.7e-05;
Matches 77; Conservative 82; Mismatches 136; Indels 53; Gaps 12;
```

```
Qy      5 EERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSD 58
Db      98 EERRQKRLQEALERQKEDPTTDSLSVPSRRMQNNSAENETAEGBEGESRSRGYEME 157
Qy      59 RAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSD 118
Db      158 ETEVATITSYQKNSYQADADKK-EKEKEEKEEKLKGNLGENQIKDEKIKQKPEKEV 216
Qy      119 QSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKL 147
Db      217 KNFLDRKKGFTVEVKAQNGEPMTKQKOTENAFSPRSQGRASGDKEAGAPQVEAGKRL 276
Qy      148 KLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKEKLQEQOSDLEQERRAKE 205
Db      277 ELRRRGETESEEFKQKQEQEALTELE-ELKKRBERRRKYLEEEOGRKQEQEADRKAR 335
Qy      206 DLEQKADTKKLERKKEHGDILADLYGRLEIPALEPSENERGY--YIPQSSLPQDN 263
Db      336 E-EERKRRLKEIERRR-----AEAAERQKMPEDGL-SEDKPKPCFTPKSGSLKIE 387
Qy      264 RGN-SRDSKETSIIEKTNRESITTVNVEGRDILHKGHLE-EKQDSIKP 309
Db      388 RAELPKNSVQKSGVASTHOAAVVSKDLSRLQYTNALIEGTAKSKMPKP 435
```

```
RESULT 10
US-10-878-556A-136
; Sequence 136, Application US/10878556A
; Publication No. US2005026399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 136
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/trbl_human
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:12:18 : Search time 8.91686 Seconds

(without alignments)
1262.481 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603

Sequence: 1 LQEQQRDLQKADTKKLE.....GHLEKKDGSIKPKKEDKS 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 80:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	602	99.8	1909	2	A45592
2	102	16.9	210	2	C70309
3	94.5	15.7	772	2	I50463
4	91.5	15.2	635	1	I37060
5	90.5	15.0	1085	2	S55352
6	90	14.9	756	2	T00367
7	89.5	14.8	3488	2	T34418
8	88	14.6	762	2	G88436
9	88	14.6	791	2	T24435
10	87	14.4	550	2	T32560
11	87	14.4	779	2	E54024
12	84.5	14.0	585	1	A24168
13	84.5	14.0	605	1	I37061
14	84.5	14.0	765	1	ISHUT1
15	84	13.9	913	2	T52485
16	83.5	13.8	615	2	T06108
17	83.5	13.8	789	2	T51310
18	83.5	13.8	821	2	S67087
19	83.5	13.8	1025	2	A47453
20	83.5	13.8	1706	2	T39305
21	83	13.8	483	2	S41853
22	83	13.8	684	2	S52835
23	82.5	13.7	495	1	I37062
24	82.5	13.7	1027	2	S37711
25	81.5	13.5	442	2	D70315
26	81.5	13.5	980	2	E71606
27	81.5	13.5	1829	2	T26135
28	81	13.4	543	2	T32973
29	81	13.4	1390	2	T14004

30	80.5	13.3	678	2	A54514	glutamic acid-rich
31	80.5	13.3	768	2	H54024	protein kinase (EC
32	80.5	13.3	777	2	B54024	protein kinase (EC
33	80.5	13.3	1032	2	I38510	neuronal kinesin h
34	80.5	13.3	1804	2	T34518	nestin - golden ha
35	80	13.3	452	2	D64583	hypothetical prote
36	80	13.3	729	2	T50989	hypothetical prote
37	79.5	13.2	501	2	C71948	hypothetical prote
38	79.5	13.2	816	2	S64439	hypothetical prote
39	79.5	13.2	1132	2	T43483	transactin initia
40	79.5	13.2	1203	2	S26650	DNA-binding protei
41	79.5	13.2	1840	2	T29091	transactin - chic
42	79.5	13.1	511	2	H83339	probable transcrip
43	79	13.1	630	2	S29796	hypothetical prote
44	79	13.1	721	2	S29795	hypothetical prote
45	79	13.1	783	2	A55817	cyclin-dependent k

ALIGNMENTS

RESULT 1
A45592
Liver stage antigen ISA-1 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 22-Nov-1993 #sequence revision 02-Dec-1994 #text change 09-Jul-2004
C:Accession: S24597; A45592; S29393; S34842; B45592; C45592; D45592
R:Zhu, J.; Hollingdale, M.
Submitted to the EMBL Data Library, November 1990
A:Reference number: S24597
A:Accession: S24597
A:Molecule type: DNA
A:Residues: 1-1909 <ZHU>
A:Cross-references: UNIPROT: Q25893; UNIPARC: UP10000081753; EMBL: X56203; NID: g9915; PID:
R:Zhu, J.; Hollingdale, M.R.
Mol. Biochem. Parasitol. 48, 223-226, 1991
A:Title: Structure of Plasmodium falciparum liver stage antigen-1.
A:Reference number: A45592; MUID: 92107224; PMID: 1840628
A:Accession: A45592
A:Molecule type: DNA
A:Residues: 1-195; 638-688; 1165-1215; 1590-1909 <ZH2>
A:Cross-references: UNIPARC: UP10000177378; UNIPARC: UP10000177379; UNIPARC: UP10000177377
A:Note: Sequence extracted from NCBI backbone (NCBI: 75010, NCBI: 75012, NCBI: 75014, NCBI: 75016)
R:Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Patrapotikul, J.; Beaudou
Nature 329, 164-167, 1987
A:Title: A liver-stage-specific antigen of Plasmodium falciparum characterized by gene
A:Reference number: S29393; MUID: 87315391; PMID: 3306406
A:Accession: S29393
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 323-387 <GUE1>
A:Cross-references: UNIPARC: UP1000017737C; EMBL: M28266
R:Guerin-Marchand, C.; Druilhe, P.; Galey, B.; Londono, A.; Patrapotikul, J.; Beaudou
Submitted to the EMBL Data Library, April 1992
A:Description: A liver-stage-specific antigen of Plasmodium falciparum characterized by
A:Reference number: S34842
A:Accession: S34842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 323-381, 'HKAI', 'GUE2'
A:Cross-references: UNIPARC: UP1000017737D; EMBL: M28266
A:Note: difference at carboxyl end due to frameshift error
C:Comment: This protein is found as flocculent material in the parasitophorous vacuole.
C:Superfamily: trichohyalin; calmodulin repeat homology
C:Keywords: EF hand
F:154-1629/Region: 17-residue repeats (A-K-E-K-L-Q-E-Q-Q-S-D-L-E-Q-E-R-R)
Query Match 99.8%, Score 602, DB 2, Length 1909;
Best Local Similarity 99.1%, Pred. No. 1, 1e-43;
Matches 116, Conservative 1, Mismatches 0, Indels 0, Gaps 0;

1 LQEQQRDLQKADTKKLEGGDILADLYGRLEIPALIEPSENERGYVPHOSSL 60
|||||

```

Db      1620 LQEQORDEQRKADTKNLERKKEHGDVLAEDLYGRLEIPALIELPSENERGYIPIHOSL 1679
QY      61 |PODNGNSRDSKEISITTEKNTRESITTVNTEGRDRIHKGHEKKGSIKEQEKDS 117
        |||
Db      1680 |PODNGNSRDSKEISITTEKNTRESITTVNTEGRDRIHKGHEKKGSIKEQEKDS 1736
        |||

RESULT 2
C70309
transcription regulator - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: C70309
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70309
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-210 <AOF>
A:Cross-references: UNIPROT:O6502; UNIPARC:UPI000056271; GB:AE000673; NID:92982834; PI
A:Experimental source: strain VF5
C:Genetics:
A:Gene: htn

Query Match      16.9%; Score 102; DB 2; Length 210;
Best Local Similarity 28.9%; Pred. No. 0.097;
Matches 43; Conservative 20; Mismatches 40; Indels 46; Gaps 7;

QY      3 EQQRDLQQRKAD-----TKNLERKKEHGDVLAEDLYGRLEIPALIELPSENERGYIPIH 56
        |||
Db      33 ERIGDLKPRKADVRIITAAITNKNLKLVEBSK-FREDLYRLNVNVELPPLERKEDITH 91
        |||

QY      57 QSS-----LPDNGNSRDSKEISITTEKNTRESI---TTN 88
        |||
Db      92 LVSYLEKFSKINGKVKVGVSPAMKLLEVDYFPGVRE-----LNNIHAHVVSSTS 145
        |||

QY      89 VEGRADIHGHEKKGDS-KEPEQEKD 116
        |||
Db      146 LIGPDDL-----EIRIGSKVKKKPRENE 170
        |||

RESULT 3
I50463
protein kinase - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 05-Oct-2004
C:Accession: I50463
R:Li, H.; Grenet, J.; Valentin, M.; Lahti, J.M.; Kidd, V.J.
Gene 153, 237-242, 1995
A:Title: Structure and expression of chicken protein kinase PITSLRE-encoding genes.
A:Reference number: I50463; MUID:95180725; PMID:7875595
A:Accession: I50463
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-772 <PIX>
A:Cross-references: UNIPROT:Q91013; UNIPARC:UPI00000FC347; EMBL:U16656; NID:9571458; PID
C:Genetics:
A:Gene: PITSLRE
A:Initons: 37/3; 76/2; 118/1; 164/2; 208/1; 249/1; 283/1; 329/1; 351/1; 407/3; 438/1; 47
C:Keywords: ATP
F:413-668/Domain: protein kinase homology <KIN>
F:421-429/Region: protein kinase ATP-binding motif

Query Match      15.7%; Score 94.5; DB 2; Length 772;
Best Local Similarity 29.3%; Pred. No. 1.9;
Matches 36; Conservative 20; Mismatches 48; Indels 19; Gaps 6;

QY      1 LQEQORDEQ-RKADTK--KNLERKKEHGDVLAEDLYGRLEIPALIELPSENERGYIPIHQ 57
        |||
Db      17 LQKKRKKEQBEKAEIKRMNSDDRSKDSLSE---GELRDHMETTINS-----PYR 68
        |||

```

```

QY      58 SLPDNGNSRDS-----KEISITTEKNTRESITTVNTEGRDRIHKGHEKKGSIKE 111
        |||
Db      69 REDSWEDEGEEDSLAIKPPQOMSREKETHHRKDEKREKR--HRSASAEKARVXEK 126
        |||

QY      112 QKE 114
        |||
Db      127 ERE 129
        |||

RESULT 4
I137060
involucrin L - gorilla
C:Species: Gorilla gorilla (gorilla)
C:Date: 16-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: I137060
R:Teuner, J.; Green, H.
Proc. Natl. Acad. Sci. U.S.A. 86, 1283-1286, 1989
A:Title: Divergent evolution of part of the involucrin gene in the hominoids: Unique int
A:Reference number: I137060; MUID:89145226; PMID:2919176
A:Accession: I137060
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-635 <RES>
A:Cross-references: UNIPROT:Q28423; UNIPARC:UPI0000087628; GB:M23603; GB:J04499; NID:934
C:Comment: During the terminal differentiation of keratinocytes, this protein from the c
linked envelope under the plasma membrane.
C:Superfamily: involucrin
C:Keywords: cornified cell envelope; duplication; epidermis; tandem repeat
F:153-590/Region: 10-residue repeats (Q-E-G-Q-[PLV]-[KE]-[LH]-[PL]-E-Q)

Query Match      15.2%; Score 91.5; DB 1; Length 635;
Best Local Similarity 26.1%; Pred. No. 2.8;
Matches 36; Conservative 29; Mismatches 42; Indels 31; Gaps 8;

QY      2 EQQRDLQQRKADTKNLERKKEHGDVLAEDLYGRLEIP-----AIELP 45
        |||
Db      149 KQQLLELPQQRGELKHEQK--GQLELPEQEQEGQLLEPQEQEGQLKHEQKQGLLEP 206
        |||

QY      46 SENERGYIPIHOS-----SLPDNGNSR--DSKEISITTEKNTRESITTVNTEGRDRIHKG 99
        |||
Db      207 EQEQQLLEPQEQEGQLLEPQEQEGQLKHEQEQEGQLLEPQEQEGQLKHEQ--KQ 262
        |||

QY      100 LE-EKKDGSIK-PEKE 114
        |||
Db      263 LELPEQEQEGQLLEPQEQ 280
        |||

RESULT 5
S55352
IFH1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein L8083.9; protein YLR223c; RRP3 protein
C:Species: Saccharomyces cerevisiae
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S55352; S51446; S47477
R:Cherel, I.; Thuriaux, P.
Yeast 11, 261-270, 1995
A:Title: The IFH1 gene product interacts with a fork head protein in Saccharomyces cere
A:Reference number: S55352; MUID:95304839; PMID:7785326
A:Accession: S55352
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1085 <CHE>
A:Cross-references: UNIPROT:P39520; UNIPARC:UPI000012D3D6; EMBL:Z29488; NID:9531491; PI
R:Hallsworth, K.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 8083.
A:Reference number: S51443
A:Accession: S51446
A:Molecule type: DNA
A:Residues: 1-1085 <HAL>
A:Cross-references: UNIPARC:UPI000012D3D6; EMBL:U19027; NID:9609363; PID:9609372; MIPS:
C:Genetics:

```


F;428-436/Region: protein kinase ATP-binding motif

Reimer, J.; Green, H.

Proc. Natl. Acad. Sci. U.S.A. 86, 1283-1286, 1989

A/Title: Divergent evolution of part of the involucrin gene in the hominoids: Unique int

A/Reference number: 137060; MUID:89145226; PMID:2919176

A/Accession: 137061

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-605 <RES>

A/Cross-references: UNIPROT:Q28424; UNIPARC:UPI000008710D; GB:M23604; GB:J04499; NID:934

C/Comment: During the terminal differentiation of keratinocytes, this protein from the c

linked envelope under the plasma membrane.

C/Superfamily: Involucrin

C/Keywords: cornified cell envelope; duplication; epidermis; tandem repeat

F153-561/Region: 10-residue repeats (Q-E-G-Q-[PLV]-[KE]-[LH]-[PL]-E-Q)

Query Match 14.0%; Score 84.5; DB 1; Length 605;

Best Local Similarity 23.5%; Pred. No. 11;

Matches 32; Conservative 25; Mismatches 32; Indels 47; Gaps 7;

QY 2 QEQORDLEQRKATTKKLERKKEHGD--LAEDLYGRLEP-----AIELP 45

DB 149 KEQLLEPREGQQLKLEQK--GQLEPREGQQLLEPREGQQLKLEQKQQLP 206

QY 46 SENERGYIYPHQS-----SLPDNRGNSRDSKEISIEKTRRESITTVNVEGRDINKHGLE 101

DB 207 EQEGQQLLEPREGQQLKLEQK-----GQLEPREGQQLKLEQK-----QEQQL 244

QY 102 --EKQDGSIK--PEQKE 114

DB 245 LPEQEGQQLLEPREGQ 260

RESULT 14

ISHUT1

DNA topoisomerase (EC 5.99.1.2) - human

N/Alternate names: nicking-closing enzyme; omega-protein; relaxing enzyme; swivelase; ty

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004

C/Accession: A30887; A40008; S13821; S02397; S40643; A34422

R/D/Arpa, P.; Machlin, P.S.; Reiter, H.; Rochfield, N.F.; Cleveland, D.W.; Earnshaw,

Proc. Natl. Acad. Sci. U.S.A. 85, 2543-2547, 1988

A/Title: CDNA cloning of human DNA topoisomerase I: catalytic activity of a 67.7-kDa cat

A/Reference number: A30887; MUID:88190108; PMID:2833744

A/Accession: A30887

A/Molecule type: mRNA

A/Residues: 1-765 <DAR>

A/Cross-references: UNIPROT:P11387; UNIPARC:UPI00001421A9; GB:J03250; NID:G339805; PIDN:

R/Kunze, N.; Yang, G.; Doelberg, M.; Sundarp, R.; Knippers, R.; Richter, A.

J. Biol. Chem. 266, 9610-9616, 1991

A/Title: Structure of the human type I DNA topoisomerase gene.

A/Reference number: A40008; MUID:91236733; PMID:1851751

A/Accession: A40008

A/Molecule type: DNA

A/Residues: 1-144, 'A', 146-553, 'E', 555-765 <KUN>

A/Cross-references: UNIPARC:UPI000001730AD; GB:M60688; GB:M60689; GB:M60690; GB:M60691; C

M60702; GB:M60703; GB:M60704; GB:M60705; GB:M60706

R/Kunze, N.; Klein, M.; Richter, A.; Knippers, R.

Eur. J. Biochem. 194, 323-330, 1990

A/Title: Structural characterization of the human DNA topoisomerase I gene promoter.

A/Reference number: S13821; MUID:91099302; PMID:2176592

A/Accession: S13821

A/Molecule type: DNA

A/Residues: 1-20 <KUN>

A/Cross-references: UNIPARC:UPI00001730AE; EMBL:X52601

R/Oddou, P.; Schmidt, U.; Knippers, R.; Richter, A.

Eur. J. Biochem. 177, 523-529, 1988

A/Title: Monoclonal antibodies neutralizing mammalian DNA topoisomerase I activity.

A/Reference number: S02397; MUID:89064806; PMID:2461859

A/Accession: S02397

A/Molecule type: mRNA

A/Residues: 344-765 <ODD>

A/Cross-references: UNIPARC:UPI00001730AF; GB:M60657

R/Tamura, H.O.; Kohchi, C.; Yamada, R.; Ikeda, T.; Koiwai, O.; Patterson, E.; Keene, J.D

Nucleic Acids Res. 19, 69-75, 1991

A/Title: Molecular cloning of a cDNA of a camptothecin-resistant human DNA topoisomerases

A/Reference number: S40643; MUID:91187651; PMID:1849260

A/Accession: S40643

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 523-543; 573-582, 'D', 584-593 <TAU>

A/Cross-references: UNIPARC:UPI00001730B0; UNIPARC:UPI00001730B1

R/Mull, G.G.; Jimenez, S.A.; Riggs, E.; Ziemnicka-Kocula, D.

Proc. Natl. Acad. Sci. U.S.A. 86, 8492-8496, 1989

A/Title: Determination of an epitope of the diffuse systemic sclerosis marker antigen D

ity in systemic sclerosis.

A/Reference number: A34422; MUID:90046823; PMID:2479024

A/Accession: A34422

A/Molecule type: mRNA

A/Residues: 657-765 <MAU>

A/Cross-references: UNIPARC:UPI000016B0FE; GB:M27913; NID:G339807; PIDN:AAA61208.1; PID

C/Comment: Type I DNA topoisomerase catalyzes the ATP-independent transient breakage of

in another, followed by rejoining. This reaction will lead to the conversion of one to

C/Genetics:

A/Gene: GDB:TOP1

A/Cross-references: GDB:120444; OMIM:126420

A/Map position: 20q12-20q13.1

A/Intons: 11/3; 20/1

A/Note: the list of introns is incomplete

C/Superfamily: eukaryotic type I DNA topoisomerase

C/Keywords: DNA binding; DNA replication; heterotetramer; isomerase

F1723/Active site: Tyr #status predicted

Query Match 14.0%; Score 84.5; DB 1; Length 765;

Best Local Similarity 25.6%; Pred. No. 14;

Matches 31; Conservative 22; Mismatches 47; Indels 21; Gaps 5;

QY 3 EQQRLDQRKATTKKLERKKEHGD--LAEDLYGRLEPAILPSENERGYIYPHOSL 60

DB 75 EKHKDKHQRDKRKRKREKRAAGDAKIKKEKENGSSPQIDBEPED--GYVPPKEDI 133

QY 61 -----PDNRGNSRDSKEISIEKTRRESITTVNVEGRDINKHGLEKQKQSIKPEQED 115

DB 134 KPLKRPDEDVDVYKPKIK--TDTYKKE-----KKLEBEDGKLKPKNKD 180

QY 116 K 116

DB 181 K 181

RESULT 15

T52485

neurofilament protein NF-M(2), middle molecular weight (imported) - African clawed frog

C/Species: Xenopus laevis (African clawed frog)

C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004

C/Accession: T52485

R/Gervasi, C.; Szabo, B.G.

Brain Res. Mol. Brain Res. 48, 229-242, 1997

A/Title: Sequence and expression pattern of two forms of the middle molecular weight ne

A/Reference number: Z26090

A/Accession: T52485

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-913 <GER>

A/Cross-references: UNIPROT:O13099; UNIPARC:UPI00000FCFAD; EMBL:U85970; PIDN:AAB53390.1

Query Match 13.9%; Score 84; DB 2; Length 913;

Best Local Similarity 26.5%; Pred. No. 19;

Matches 30; Conservative 21; Mismatches 50; Indels 12; Gaps 4;

QY 2 QEQORDLEQRKATTKKLERKKEHGDLAEDLYGRLEPAILPSENERGYIYPHOSLP 61

DB 550 ENEEGGQGEEDVYKKEEKEGGERDAE---VEETKAEDEBEDEKQ-----EDBAE 600

QY 62 PDNRGNSRDSKEISIEKTRRESITTVNVEGRDINKHGLEKQKQSIKPEQKE 114

DB 601 KEKGEEBEKEVVAABDE--GGEDEEGGEGEKEDKG--BEKEGEABAEAE 650

Search completed: December 7, 2005, 07:21:23
Job time : 10.9169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:11:08 ; Search time 37.0165 Seconds
(without alignments)
2229.881 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LQEQQRDLQKADTKKNL.....GHLEKKDGSIKPEQKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	603	100.0	493	Q25886_PLAFA	Q25886 plasmodium
2	602	99.8	1909	Q25893_PLAFA	Q25893 plasmodium
3	597	99.0	1596	Q81J44_PLAF7	Q81J44 plasmodium
4	554	91.9	280	Q25844_PLAFA	Q25844 plasmodium
5	554	91.9	280	Q25849_PLAFA	Q25849 plasmodium
6	554	91.9	280	Q25850_PLAFA	Q25850 plasmodium
7	553	91.7	280	Q25843_PLAFA	Q25843 plasmodium
8	553	91.7	280	Q25846_PLAFA	Q25846 plasmodium
9	553	91.7	280	Q25847_PLAFA	Q25847 plasmodium
10	553	91.7	280	Q25851_PLAFA	Q25851 plasmodium
11	550	91.2	280	Q25848_PLAFA	Q25848 plasmodium
12	550	91.2	280	Q25852_PLAFA	Q25852 plasmodium
13	549	91.0	280	Q25845_PLAFA	Q25845 plasmodium
14	549	91.0	280	Q25853_PLAFA	Q25853 plasmodium
15	549	91.0	280	Q25854_PLAFA	Q25854 plasmodium
16	549	91.0	280	Q25855_PLAFA	Q25855 plasmodium
17	549	91.0	280	Q25888_PLAFA	Q25888 plasmodium
18	549	91.0	280	Q25889_PLAFA	Q25889 plasmodium
19	549	91.0	280	Q25900_PLAFA	Q25900 plasmodium
20	549	91.0	280	Q27243_PLAFA	Q27243 plasmodium
21	469	77.8	264	Q9GTXS_PLAFA	Q9GTXS plasmodium
22	224	37.1	42	Q8WZK9_PLAFA	Q8WZK9 plasmodium
23	102	16.9	210	Q66502_AOUAB	Q66502 equifex ae
24	98.5	16.3	924	Q81445_PLAF7	Q81445 plasmodium
25	97	16.1	1130	Q59YV6_CANAL	Q59YV6 candida alb
26	94.5	15.7	772	Q91013_CHICK	Q91013 gallus gall
27	94	15.6	245	Q51017_ENTHI	Q51017 entamoeba h
28	94	15.6	1813	Q51393_ENTHI	Q51393 entamoeba h
29	93.5	15.5	3812	Q4RLC8_TERNI	Q4RLC8 tetraodon n
30	92	15.3	603	Q9AHN3_PASMU	Q9AHN3 pasteurella
31	92	15.3	783	Q4T6N6_TERNI	Q4T6N6 tetraodon n

32	91.5	15.2	635	Q28423_9PRIM	Q28423 gorilla gor
33	91	15.1	294	Q095S6_CLORE	Q095S6 clostridium
34	90.5	15.0	1085	1 IFH1 YEAST	P39520 saccharomyc
35	90	14.9	225	Q725E9_HUMAN	Q725E9 homo sapien
36	90	14.9	756	1 RPTP3_HUMAN	Q75154 homo sapien
37	90	14.9	756	Q4VXV7_HUMAN	Q4VXV7 homo sapien
38	89.5	14.8	295	Q6BXK9_DEBIA	Q6BXK9 debaryomyc
39	89.5	14.8	485	Q75EU2_ASHCO	Q75EU2 ashya goss
40	89.5	14.8	18519	Q81SF7_CAEEL	Q81SF7 caenorhabdi
41	89.5	14.8	18519	Q81SF7_CAEEL	Q81SF7 caenorhabdi
42	88.5	14.7	752	Q4Z6S0_PLABE	Q4Z6S0 plasmodium
43	88.5	14.7	807	Q5TM60_MACMU	Q5TM60 macaca mula
44	88.5	14.7	847	Q4FPL4_9RIKX	Q4FPL4 candidatus
45	88.5	14.7	1405	Q4Z3Q7_PLABE	Q4Z3Q7 plasmodium

ALIGNMENTS

```
RESULT 1
Q25886_PLAFA PRELIMINARY; PRT; 493 AA.
ID Q25886_PLAFA PRELIMINARY; PRT; 493 AA.
AC Q25886;
DT 01-NOV-1996 (TREMUREL. 01, Created)
DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
DT 01-OCT-2002 (TREMUREL. 22, Last annotation update)
DE Liver stage antigen-1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5833;
RX MEDLINE=9426724; Pubmed=7515922;
RA Fidock D.A., Gras-Masse H., Lepers J., Brahimi K., Bernhamed L.,
RA Meis J.F., Langley G., Rousillon C., Londono A., Raharimalala L.,
RA Meis J.F., Langley G., Rousillon C., Tartar A., Druilhe P.;
RT "Plasmodium falciparum liver stage antigen-1 is well conserved and
RT contains potent B and T cell determinants."
RT J. Immunol. 153:190-204 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Fidock D.A., Gras-Masse H., Lepers J., Brahimi K., Bernhamed L.,
RA Weillou S., Guerin-Marchand C., Londono A., Raharimalala L.,
RA Meis J.F., Langley G., Rousillon C., Tartar A., Druilhe P.;
RT "The plasmodium falciparum liver stage antigen LSA-1 is well conserved
RT and harbors major B- and T-cell epitopes."
RL J. Immunol. 0:0-0 (1994).
DR EMBL, Z30320; CAA82975.1; -; Genomic_DNA.
FT NON_TER
SQ SEQUENCE 493 AA; 59062 MW; FE252CC44F46996E CRC64;

Query Match 100.0%; Score 603; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.7e-41;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEQQRDLQKADTKKNLKKKHGDIILADLYGRLEIPALIEPSNENGYIIPHOSSL 60
Db 204 LQEQQRDLQKADTKKNLKKKHGDIILADLYGRLEIPALIEPSNENGYIIPHOSSL 263

QY 61 PQRKNGSRBSKESITTEKTNRESITTNVGRPDHKGHEKKDGSIKPEQKEDKS 117
Db 264 PQRKNGSRBSKESITTEKTNRESITTNVGRPDHKGHEKKDGSIKPEQKEDKS 320

RESULT 2
Q25893_PLAFA PRELIMINARY; PRT; 1909 AA.
ID Q25893_PLAFA PRELIMINARY; PRT; 1909 AA.
AC Q25893;
DT 01-NOV-1996 (TREMUREL. 01, Created)
DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
DT 01-JUN-2003 (TREMUREL. 24, Last annotation update)
DE Liver stage antigen.
```

GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NF54;
RX MEDLINE=92107224; PubMed=1840628; DOI=10.1016/0166-6851(91)90117-0;
RT "Zhu J., Hollingdale M.R.; Plasmodium falciparum liver stage antigen-1.";
RT "Structure of Plasmodium falciparum liver stage antigen-1.";
RL Mol. Biochem. Parasitol. 48:223-228(1991).
DR EMBL; X56203; CAA39663.1; -; Genomic_DNA.
DR PIR; S24597; A45592.
DR HSSP; 015813; 1D7M.
SQ SEQUENCE 1909 AA; 230153 MW; 8B861A524C797217 CRC64;

Query Match 99.8%; Score 602; DB 2; Length 1909;
Best Local Similarity 99.1%; Pred. No. 9.7e-41;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSL 60
DB 1620 LQEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSL 1679
QY 61 PQQNRGNSRDSKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 117
DB 1680 PQQNRGNSRDSKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 1736

RESULT 3
Q81J44_PLAF7
ID Q81J44_PLAF7 PRELIMINARY; PRT; 1596 AA.

AC Q81J44;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Liver stage antigen, putative.
GN ORFNames=Pr10.0356;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=22255705; PubMed=12368664; DOI=10.1038/nature01097;
RX Gardner M.J., Hall N., Fung B., White O., Berriman M., Hyman R.W.,
RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow J., Suh B., Peterson J., Angiuoli S.,
RA Perera M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Sudramanlan G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.G.;
RT "Genome sequence of the human malaria parasite Plasmodium falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; A6014834; AAN3553.1; -; Genomic_DNA.
SQ SEQUENCE 1596 AA; 191691 MW; COB3A2CC0F5D60AC CRC64;

Query Match 99.0%; Score 597; DB 2; Length 1596;
Best Local Similarity 98.3%; Pred. No. 2e-40;
Matches 115; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LQEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSL 60
DB 1307 LQEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSL 1366
QY 61 PQQNRGNSRDSKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 117
DB 1367 PQQNRGNSRDSKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 1423

RESULT 4

Q25844_PLAFA
ID Q25844_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25844;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates.";
RT Isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40885; AAB59231.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32940 MW; E9708B3CFPAEA9CF CRC64;

Query Match 91.9%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RQADTKNLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSLPQDNRGNSRD 70
DB 1 RQADTKNLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSLPQDNRGNSRD 60
QY 71 SKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 117
DB 61 SKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 107

RESULT 5
Q25849_PLAFA
ID Q25849_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25849;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
RA Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates.";
RT Isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
DR EMBL; L40890; AAC1600.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32966 MW; E96D255154DEA9CF CRC64;

Query Match 91.9%; Score 554; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RQADTKNLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSLPQDNRGNSRD 70
DB 1 RQADTKNLERKKEHGDILAEDLYGRLEIPALPSENERGYIIPHOSSLPQDNRGNSRD 60
QY 71 SKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 117
DB 61 SKEISITTEKTNRESITTVNVEGRDIIHGHLERKKGSIKPEQKEDKS 107

```
RESULT 6
Q25850_PLAFA
ID Q25850_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25850;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
EMBL; L40891; AAC41601.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32989 MW; E96F812CFABEBD8 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 9.8e-38;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 70
DB 1 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEEKDGSIKPEQKEDKS 117
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEEKDGSIKPEQKEDKS 107

RESULT 7
Q25843_PLAFA
ID Q25843_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25843;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
EMBL; L40884; AAB59230.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32926 MW; 467080F3FAEADJ3 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 70
DB 1 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60
```

```
QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEEKDGSIKPEQKEDKS 117
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEEKDGSIKPEQKEDKS 107

RESULT 8
Q25846_PLAFA
ID Q25846_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25846;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
EMBL; L40887; AAC41597.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32882 MW; 46763641C81AFC33 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 70
DB 1 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEEKDGSIKPEQKEDKS 117
DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEEKDGSIKPEQKEDKS 107

RESULT 9
Q25847_PLAFA
ID Q25847_PLAFA PRELIMINARY; PRT; 280 AA.
AC Q25847;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-
specific antigen-1 (LSA-1) of Plasmodium falciparum from field
isolates.";
RL Mol. Biochem. Parasitol. 71:291-294(1995).
EMBL; L40886; AAC41598.1; -, Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 280 AA; 32944 MW; 467085E23BABBC27 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;
Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKKLERKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSLPODNRGNSRD 70
```

Db 1 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHOSLPODNRGNSRD 60

Qy 71 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEOKEDKS 117

Db 61 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEOKEDKS 107

RESULT 10

Q25851_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25851;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxId=5833;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RT Mol. Biochem. Parasitol. 71:291-294(1995).
RL EMBL; L40893; AAC1603.1; -; Genomic_DNA.
RT NON_TER 1
SQ SEQUENCE 280 AA; 32927 MW; 48706EFDCCF40ADD3 CRC64;

Query Match 91.7%; Score 553; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 1.2e-37;

Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHOSLPODNRGNSRD 70

Db 1 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHOSLPODNRGNSRD 60

Qy 71 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEOKEDKS 117

Db 61 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEOKEDKS 107

RESULT 11

Q25848_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25848;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxId=5833;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RT Mol. Biochem. Parasitol. 71:291-294(1995).
RL EMBL; L40889; AAC1599.1; -; Genomic_DNA.
RT NON_TER 1
SQ SEQUENCE 280 AA; 32939 MW; E77080385FBA0D61 CRC64;

Query Match 91.2%; Score 550; DB 2; Length 280;
Best Local Similarity 99.1%; Pred. No. 2.1e-37;

Matches 106; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHOSLPODNRGNSRD 70

Db 1 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHOSLPODNRGNSRD 60

Qy 71 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEOKEDKS 117

Db 61 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEOKEDKS 107

RESULT 12

Q25852_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25852;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxId=5833;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RT Mol. Biochem. Parasitol. 71:291-294(1995).
RL EMBL; L40834; AAB59232.1; -; Genomic_DNA.
RT NON_TER 1
SQ SEQUENCE 280 AA; 32927 MW; E6648F85FBBA245 CRC64;

Query Match 91.2%; Score 550; DB 2; Length 280;
Best Local Similarity 98.1%; Pred. No. 2.1e-37;

Matches 105; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHOSLPODNRGNSRD 70

Db 1 RKADTKNLERKKEGDIADLYGRLEIPALIELPSENERGYIIPHOSLPODNRGNSRD 60

Qy 71 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEOKEDKS 117

Db 61 SKEISIIKTNRESITTNVEGRDIIHKHLEKKDGSIKPEOKEDKS 107

RESULT 13

Q25845_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25845;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Liver stage-specific antigen 1 (Fragment).
GN Name=LSA-1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxId=5833;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=96065765; PubMed=7477115; DOI=10.1016/0166-6851(95)00069-D;
RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,
Hawley W.A., Collins W.E., Lal A.A.;
RT "Sequence variations in the non-repetitive regions of the liver stage-specific antigen-1 (LSA-1) of Plasmodium falciparum from field isolates."
RT Mol. Biochem. Parasitol. 71:291-294(1995).
RL EMBL; L40886; AAB59232.1; -; Genomic_DNA.
RT NON_TER 1
SQ SEQUENCE 280 AA; 32908 MW; FDEE929F34DCA7E9 CRC64;

Query Match 91.0%; Score 549; DB 2; Length 280;
 Best Local Similarity 98.1%; Pred. No. 2.5e-37;
 Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

11 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 70
 1 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
 DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107

RESULT 14
 Q25853_PLAFA
 ID Q25853_PLAFA PRELIMINARY; PRT; 280 AA.
 AC Q25853

DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Liver stage-specific antigen 1 (Fragment).

GN Name=LSA-1;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_Taxid=5833;
 RN 11

RP NUCLEOTIDE SEQUENCE.

RA MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;

RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,

RA Hawley W.A., Collins W.E., Lal A.A.,

RT "Sequence variations in the non-repetitive regions of the liver stage-

RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field

RT isolates.";

RT Mol. Biochem. Parasitol. 71:291-294(1995).

DR EMBL: I40835; AAB59234.1; -, Genomic_DNA.

FT NON_TER

SEQUENCE 280 AA; 32938 MW; 46751C45F4DCBD33 CRC64;

Query Match 91.0%; Score 549; DB 2; Length 280;
 Best Local Similarity 98.1%; Pred. No. 2.5e-37;
 Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 70
 DB 1 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
 DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107

RESULT 15

Q25854_PLAFA

ID Q25854_PLAFA PRELIMINARY; PRT; 280 AA.

AC Q25854

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Liver stage-specific antigen 1 (Fragment).

GN Name=LSA-1;

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_Taxid=5833;

RN 11

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=96065765; Pubmed=7477115; DOI=10.1016/0166-6851(95)00069-D;

RA Yang C., Shi Y.P., Udhayakumar V., Alpers M.P., Povoa M.M.,

RA Hawley W.A., Collins W.E., Lal A.A.,

RT "Sequence variations in the non-repetitive regions of the liver stage-

RT specific antigen-1 (LSA-1) of Plasmodium falciparum from field

RT isolates.";

RT Mol. Biochem. Parasitol. 71:291-294(1995).

DR EMBL: I40836; AAC1604.1; -, Genomic_DNA.

FT NON_TER 1 1
 SEQUENCE 280 AA; 32966 MW; 61851C5ADE1A211 CRC64;

Query Match 91.0%; Score 549; DB 2; Length 280;
 Best Local Similarity 98.1%; Pred. No. 2.5e-37;
 Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 70
 DB 1 RKADTKKNLERKKEHGDVLAEDLYGRLEIPALIEPSENERGYIIPHOSSLPODNRGNSRD 60

QY 71 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
 DB 61 SKEISIIKTNRESITTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 107

Search completed: December 7, 2005, 07:20:44
 Job time : 39.0185 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 06:57:58 ; Search time 37.8291 Seconds
(without alignment)
1358.936 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LQEQQRDLQKADTKKLE.....GLLEKKKGSIRPEKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_21.*
1: geneseqp19808.*
2: geneseqp19908.*
3: geneseqp20008.*
4: geneseqp20018.*
5: geneseqp20028.*
6: geneseqp20038.*
7: geneseqp20048.*
8: geneseqp20058.*
9: geneseqp20058.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603	100.0	117	2	AAR26937 P.falcipa
2	603	100.0	316	2	AAR26941 P.falcipa
3	603	100.0	493	2	AAR26944 P.falcipa
4	602	99.8	456	8	ADO21942 LSA-NRC(H
5	602	99.8	457	8	ADO21920 LSA-NRC(H
6	172	28.5	34	8	ADO21926 Malaria p
7	103	17.1	20	8	ADO21925 Malaria p
8	94	15.6	452	9	Ad240617 PfCTL/HTL
9	94	15.6	456	7	Ad240617 Multi-epi
10	94	15.6	456	8	Ado24115 Epigene c
11	90.5	15.0	1085	2	AAR95607 RRP3 telo
12	90.5	15.0	1085	7	ADK62464 Disease t
13	89.5	14.8	231	1	AAG98701 Human cel
14	89.5	14.8	232	4	AAG98700 Human cel
15	89.5	14.8	236	4	AAG98699 Human cel
16	89.5	14.8	257	4	AAG98698 Human cel
17	89.5	14.8	272	4	AAG98697 Human cel
18	87.5	14.5	271	5	ABB04736 Human cyl
19	87.5	14.5	282	3	AAV58635 Protein r
20	86	14.3	550	8	ADN23342 Bacterial
21	86	14.3	447	4	AAG82903 S. epider
22	85.5	14.2	450	4	AAG82906 S. epider
23	85.5	14.2	684	8	Ad116245 Human nuc
24	85.5	14.2	765	6	ABM35013 Cancer ba

25	85.5	14.2	765	6	ADB20866 MKPI baee
26	85.5	14.2	765	6	ABB82815 Human top
27	85.5	14.2	765	7	ADB87955 Human UGT
28	85.5	14.2	765	7	ADB96938 Human MDR
29	85.5	14.2	765	7	ADB92129 Human MDR
30	85.5	14.2	765	7	ADJ69343 Human hea
31	85.5	14.2	765	8	ADJ66581 DNA topol
32	85.5	14.2	765	8	ADL82893 Human PRO
33	85.5	14.2	765	8	ADP56068 Human PRO
34	85	14.1	15	2	AAW85415 Helper T-
35	85	14.1	15	2	AAW85172 Helper T-
36	85	14.1	15	7	Adw36112 HLA bindi
37	85	14.1	15	7	ADW35943 HLA bindi
38	84.5	14.0	150	5	ABB08082 Human top
39	84.5	14.0	242	4	ABG02495 Novel hum
40	84.5	14.0	300	5	ABB08081 Human top
41	84.5	14.0	323	8	ADN04692 Human top
42	84.5	14.0	585	8	ADJ75618 Marker ge
43	84.5	14.0	585	9	ADX06678 Cyclin-de
44	84.5	14.0	765	1	ADP92275 Human top
45	84.5	14.0	765	2	AAR75915 Variant h

ALIGNMENTS

RESULT 1
AAR26937
ID AAR26937 standard; peptide; 117 AA.
XX
XX AAR26937;
AC
XX 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
XX P.falciparum liver-stage specific antigen non-repeat region.
DE
XX
XX
XX Malaria; LSA-NR; hepatocyte; sporozoite; plasmodium DG 536; T-cell epitope; paludism.
KM
XX
XX Plasmodium falciparum.
OS
XX
XX WO9213884-A1.
PN
XX
XX 20-AUG-1992.
PD
XX
XX 05-FEB-1992; 92WO-FR000104.
PF
XX
XX 05-FEB-1991; 91FR-00001286.
PR
XX
XX (INSP) INST PASTEUR.
XX
XX Guerinmarchand C, Drullhe P;
PI
XX WPI; 1992-299985/36.
PT
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX
XX
XX Claim 4; Page 55; 81pp; French.
XX
XX A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
CC gill was used to transform E.coli. The expression library was screened
CC with human antisera against antigens of all stages of P. falciparum. The
CC library was rescreened with antibodies affinity-purified on a clone
CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic LSA
CC epitope. The clone with the largest insert encoded LSA-R-NR containing a
CC 12-repeat region (LSA-R) followed by a non-repeat region (LSA-NR). The
CC sequence given here is a preferred polypeptide of the invention carrying
CC a T cell epitope typical of liver-stage P.falciparum. See AAR26919-
CC R26935, AAR26937 and AAR26939. (Updated on 25-MAR-2003 to correct PN
CC field.)

```
XX Sequence 117 AA;
SO Query Match 100.0%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 8,36-57;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEQORDLEQKADTKKQLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSL 60
DB 1 LOEQORDLEQKADTKKQLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSL 60

OY 61 PDNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
DB 61 PDNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117

RESULT 2
AAR26941 standard; protein; 316 AA.
XX ID AAR26941
XX AC AAR26941;
XX AC AAR26941;
XX DT 25-MAR-2003 (revised)
XX DT 08-FEB-1993 (first entry)
XX DE P.falciparum LSA-R-NR protein.
XX KW Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
XX KW paludism; liver stage-specific antigen.
XX OS Plasmodium falciparum.
XX FH Key Location/Qualifiers
XX FT 1..209
XX FT /label= repeat_region
XX FT /note= "contains 12 x 17mer repeats"
XX FT 210..316
XX FT Region /label= non-repeat_region
XX FT
XX PN WO9213884-A1.
XX PD 20-AUG-1992.
XX PD
XX PF 05-FEB-1992; 92WO-FR000104.
XX PR 05-FEB-1991; 91FR-00001286.
XX PA (INSP ) INST PASTEUR.
XX PI Guerimarchand C, Drulhe P;
XX DR WPI, 1992-299985/36.
XX DR N-PSDB; AAQ28115.
XX PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX PT vaccination against, treatment of and diagnosis of malaria.
XX PS Disclosure; Fig 1; 81pp; French.
XX CC A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
XX CC gtl1 was used to transform E.coli. The expression library was screened
XX CC with human antisera against antigens of all stages of P.falciparum. The
XX CC library was rescreened with antibodies affinity-purified on a clone
XX CC which was able to recognise antibodies specific to the hepatic phase.
XX CC About 40 clones were detected which produced a characteristic LSA
XX CC epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
XX CC containing a 12-repeat region followed by a non-repeat region. Preferred
XX CC antigenic polypeptides of the invention are derived from the amino acid
XX CC sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 316 AA;
Query Match 100.0%; Score 603; DB 2; Length 316;
```

```
Best Local Similarity 100.0%; Pred. No. 3,1e-56;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEQORDLEQKADTKKQLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSL 60
DB 200 LOEQORDLEQKADTKKQLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSL 259

OY 61 PDNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
DB 260 PDNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 316

RESULT 3
AAR26944 standard; protein; 493 AA.
XX ID AAR26944
XX AC AAR26944;
XX AC AAR26944;
XX DT 25-MAR-2003 (revised)
XX DT 08-FEB-1993 (first entry)
XX DE P.falciparum LSA gene C-terminal region.
XX KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
XX KW paludism; liver stage-specific antigen.
XX OS Plasmodium falciparum.
XX FH Key Location/Qualifiers
XX FT 13..213
XX FT /label= repetitive_region
XX FT 214..493
XX FT Region /label= non-repetitive_region
XX FT
XX PN WO9213884-A1.
XX PD 20-AUG-1992.
XX PD
XX PF 05-FEB-1992; 92WO-FR000104.
XX PR 05-FEB-1991; 91FR-00001286.
XX PA (INSP ) INST PASTEUR.
XX PI Guerimarchand C, Drulhe P;
XX DR WPI, 1992-299985/36.
XX DR N-PSDB; AAQ28119.
XX PT Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX PT vaccination against, treatment of and diagnosis of malaria.
XX PS Claim 2; Fig 8-10; 81pp; French.
XX CC The 3' part of the P.falciparum liver-stage specific antigen (LSA) gene
XX CC codes for a polypeptide sequence which carries a T cell epitope
XX CC characteristic of a protein produced in hepatocytes infected with
XX CC P.falciparum. The polypeptide can be used in the preparation of vaccines
XX CC against malaria. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 493 AA;
Query Match 100.0%; Score 603; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 5,5e-56;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LOEQORDLEQKADTKKQLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSL 60
DB 204 LOEQORDLEQKADTKKQLERKKEHGDIADLYGRLEIPALIEPSENERGYIIPHOSSL 263

OY 61 PDNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 117
DB 264 PDNRGNSRDSKEISIIIEKTNRESITTTNVEGRDIIHKHGLEKKDGSIKPEOKEDKS 320
```

RESULT 4
ADO21942
ID ADO21942 standard; protein; 456 AA.
XX
XX ADO21942;
AC
DT 12-AUG-2004 (first entry)
XX
XX LSA-NRC(H) construct protein derived from Malaria parasite LSA-1.
DE
XX LSA-NRC, liver stage antigen; LSA-1; T-cell; B-cell epitope;
XX parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KW malaria parasite P. falciparum; LSA-NRC(H) construct.
XX
XX Plasmodium falciparum.
OS Synthetic.
XX
XX WO2004044167-A2.
XX
XX 27-MAY-2004.
XX
XX 12-NOV-2003; 2003WO-US036011.
XX
XX 12-NOV-2002; 2002US-0425719P.
XX
XX (REED-) REED ARMY INST RES WALTER.
XX
XX Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barbosa A;
XX
XX WPI; 2004-420309/39.
XX
XX N-PSDB; ADO21941.
XX
XX Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
XX Claim 4; SEQ ID NO 26; 90pp; English.
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
XX
XX Sequence 456 AA;
SQ
Query Match 99.8%; Score 602; DB 8; Length 456;
Best Local Similarity 99.1%; Pred. No. 6,4e-56;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQEQORDLEQRKADTKNLERKKEHGDIIAEDLYGRLEIPALPSENENRGYYIPHOSSL 60
DB 155 LQEQORDLEQRKADTKNLERKKEHGDVLAEDLYGRLEIPALPSENENRGYYIPHOSSL 214
QY 61 PQQNRGNSRDSKEISIIETKTNRESITTNVGRDIIHKGHLEKKDGSIKPEQKEDKS 117
DB 215 PQQNRGNSRDSKEISIIETKTNRESITTNVGRDIIHKGHLEKKDGSIKPEQKEDKS 271

DT 12-AUG-2004 (first entry)
XX
XX LSA-NRC(H)Mut construct protein derived from Malaria parasite LSA-1.
DE
XX LSA-NRC, liver stage antigen; LSA-1; T-cell; B-cell epitope;
XX parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KW malaria parasite P. falciparum; LSA-NRC(H)Mut construct; mutant; mtein.
XX
XX Plasmodium falciparum.
OS Synthetic.
XX
XX WO2004044167-A2.
XX
XX 27-MAY-2004.
XX
XX 12-NOV-2003; 2003WO-US036011.
XX
XX 12-NOV-2002; 2002US-0425719P.
XX
XX (REED-) REED ARMY INST RES WALTER.
XX
XX Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barbosa A;
XX
XX WPI; 2004-420309/39.
XX
XX N-PSDB; ADO21919.
XX
XX Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
XX Claim 5; SEQ ID NO 4; 90pp; English.
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct protein of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC within the T5 P. falciparum epitope and a C-terminal His6 tag.
XX
XX Sequence 457 AA;
SQ
Query Match 99.8%; Score 602; DB 8; Length 457;
Best Local Similarity 99.1%; Pred. No. 6,4e-56;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQEQORDLEQRKADTKNLERKKEHGDIIAEDLYGRLEIPALPSENENRGYYIPHOSSL 60
DB 155 LQEQORDLEQRKADTKNLERKKEHGDVLAEDLYGRLEIPALPSENENRGYYIPHOSSL 214
QY 61 PQQNRGNSRDSKEISIIETKTNRESITTNVGRDIIHKGHLEKKDGSIKPEQKEDKS 117
DB 215 PQQNRGNSRDSKEISIIETKTNRESITTNVGRDIIHKGHLEKKDGSIKPEQKEDKS 271

RESULT 6
ADO21926
ID ADO21926 standard; peptide; 34 AA.
XX
XX ADO21926;
AC
DT 12-AUG-2004 (first entry)
XX

DE Malaria parasite P. falciparum LSA-1 LSA-ter epitope peptide.
XX
XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KM parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
XX Malaria parasite P. falciparum; epitope; LSA-ter.
XX Plasmodium falciparum.
OS
PN WO200404167-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US036011.
XX
PR 12-NOV-2002; 2002US-0425719P.
XX
PA (REED-) REED ARMY INST RES WALTER.
XX
PI Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barboea A;
XX
XX WPI; 2004-420309/39.
DR
XX
PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
PS Claim 2; SEQ ID NO 10; 90pp; English.
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the malaria parasite P. falciparum LSA-1 LSA-ter epitope
CC peptide of the invention.
XX
SQ Sequence 34 AA;
XX
Query Match 28.5%; Score 172; DB 8; Length 34;
Best Local Similarity 100.0%; Pred. No. 4,4e-11;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 67 NSRDSKESITIKTNRESITTNVGRDIIHGHL 100
DB 1 NSRDSKESITIKTNRESITTNVGRDIIHGHL 34
RESULT 7
ADO21925
ID ADO21925 standard; peptide; 20 AA.
XX
XX ADO21925;
XX
DT 12-AUG-2004 (first entry)
XX
DE Malaria parasite P. falciparum liver stage antigen-1 NR epitope peptide.
XX
XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
KM parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
KW malaria parasite P. falciparum; epitope; NR.
XX
OS Plasmodium falciparum.
XX
PN WO200404167-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US036011.
XX
PR 12-NOV-2002; 2002US-0425719P.
XX

PA (REED-) REED ARMY INST RES WALTER.
XX
XX Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Barboea A;
XX
XX WPI; 2004-420309/39.
DR
XX
XX Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
XX
PS Claim 2; SEQ ID NO 9; 90pp; English.
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the malaria parasite P. falciparum LSA-1 NR epitope peptide of
CC the invention.
XX
SQ Sequence 20 AA;
XX
Query Match 17.1%; Score 103; DB 8; Length 20;
Best Local Similarity 95.0%; Pred. No. 0.00059;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 15 TKKNLERKKEHGDVLABDLY 34
DB 1 TKKNLERKKEHGDVLABDLY 20
RESULT 8
ADZ40617
ID ADZ40617 standard; protein; 452 AA.
XX
XX ADZ40617;
AC
XX
DT 30-JUN-2005 (first entry)
XX
XX pFCTT/HTL/N multi-epitope construct protein.
DE
XX multi-epitope; immune stimulation; HIV infection; anti-HIV;
KW immunostimulant; vaccine; immunogenicity.
XX
OS Synthetic.
XX
PN WO2005033265-A2.
XX
PD 14-APR-2005.
XX
PF 26-APR-2004; 2004WO-US012732.
XX
PR 25-APR-2003; 2003US-0465229P.
XX
XX (EPIM-) EPIMMUNE INC.
PA
PI Sette A, Chesnut RW, Newman MJ, Livingston BD;
PI WPI; 2005-285416/29.
DR
DR N-PSDB; ADZ40618.
XX
XX New multi-epitope polynucleotide, useful for designing vaccines that
PT provide increased immunogenicity, and for inducing immune responses
PT against HIV in an individual.
XX
PS Disclosure; Fig 181; 261pp; English.
XX
XX The invention relates to a novel multi-epitope polynucleotide. The
CC invention further comprises: a composition comprising the polynucleotide;
CC a cell comprising the polynucleotide; a method for inducing an immune
CC response against human immunodeficiency virus (HIV) in an individual.

CC comprising administering the polynucleotide, composition or cell to the
CC individual; and a method for making the polynucleotide, composition or
CC cell. The multi-epitope polynucleotide has anti-HIV and immunostimulant
CC activities. The polynucleotide is useful for designing vaccines that
CC provide increased immunogenicity, and for inducing an immune response
CC against HIV. This sequence represents an pCTL/HTL/N multi-epitope
CC construct protein of the invention.

XX Sequence 452 AA;
SQ

Query Match 15.6%; Score 94; DB 9; Length 452;
Best Local Similarity 85.7%; Pred. No. 0.33;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 40 PAIELPSENERGYIIPHOSSL 60
|
DB 119 PGPGLPSENERGYIIPHOSSL 139

RESULT 9
ADA49437
ID ADA49437 standard; protein; 456 AA.
XX
AC ADA49437;
XX
DT 20-NOV-2003 (first entry)
XX
DE Multi-epitope construct #21.
XX
KM multi-epitope; immunogenic; epitope; major histocompatibility complex;
KM MHC class I; MHC class II; junctional epitope.
XX
OS Synthetic.
OS Plasmodium falciparum.
XX
PN US2002119127-A1.
XX
PD 29-AUG-2002.
XX
PF 27-JUN-2001; 2001US-00894018.
XX
PR 28-DEC-1999; 99US-0173390P.
PR 28-DEC-2000; 2000MO-US035568.
PR 16-APR-2001; 2001US-0284221P.
XX
PA (SETT/) SETTE A.
PA (CHES/) CHESNUT R.
PA (LIIV/) LIVINGSTON B D.
PA (BAKE/) BAKER D M.
PA (NEWM/) NEWMAN M J.
PA (BROW/) BROWN D H.
XX
PI Sette A, Chesnut R, Livingston BD, Baker DM, Newman MJ, Brown DH;
XX WPI, 2003-615704/58.
XX N-PSDB; ADA49438.
XX
PT Designing multi-epitope construct having major histocompatibility complex
PT class I and II epitope nucleic acids, by selecting mixture of amino acid
PT insertions at junctions of construct to minimize junctional epitopes.
XX
XX Disclosure; Fig 18f; 78pp; English.

CC The invention relates to a method of designing multi-epitope constructs
CC comprising major histocompatibility complex (MHC) class I and II (CTL)
CC epitope nucleic acids (CEN), involves sorting CEN, introducing flanking
CC amino acid residue selected from specified amino acid residues given in
CC specification at C+1 position of CEN, introducing amino acid spacer
CC residues between two CEN, and selecting the constructs having less
CC junctional epitopes. The method is useful for designing a multi-epitope
CC construct having multiple nucleic acid. The method avoids or
CC minimizes the occurrence of junctional epitopes and maximizes the
CC immunogenicity and/or antigenicity of multi-epitope vaccines. The present

CC sequence represents the amino acid sequence of a multi-epitope construct.
XX
SQ Sequence 456 AA;
SQ

Query Match 15.6%; Score 94; DB 7; Length 456;
Best Local Similarity 85.7%; Pred. No. 0.34;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 40 PAIELPSENERGYIIPHOSSL 60
|
DB 119 PGPGLPSENERGYIIPHOSSL 139

RESULT 10
AD024115
ID AD024115 standard; protein; 456 AA.
XX
AC AD024115;
XX
DT 01-UTL-2004 (first entry)
XX
DE Epigene construct pCTL/HTL(N) protein.
XX
KM hepatotropic; virucide; anti-inflammatory; anti-HIV; cytostatic; epitope;
KM cytotoxic T lymphocyte; CTL; helper T lymphocyte; HTL; immune response;
KM hepatitis B virus; HBV; HIV; cancer; lymphoma.
XX
OS Unidentified.
XX
PN MO2004031210-A2.
XX
PD 15-APR-2004.
XX
PF 03-OCT-2003; 2003MO-US031303.
XX
PR 03-OCT-2002; 2002US-0415463P.
PR 22-OCT-2002; 2002US-0419973P.
XX
PA (EPIM-) EPIMUNE INC.
PA (GENV-) GENENCOR INT INC.
XX
PI Sette A, Chesnut R, Newman MJ, Livingston BD, Babe LM, Chen Y;
PI Deyoung LM, Huang MTF, Power SD;
XX WPI, 2004-330143/30.
XX N-PSDB; AD024116.
XX
PT New multi-epitope polynucleotides encoding cytotoxic T lymphocyte and/or
PT helper T lymphocyte epitopes, useful for inducing or stimulating an
PT antiviral or anticancer immune response, especially against hepatitis B
PT virus.
XX
XX Disclosure; SEQ ID NO 265; 401pp; English.

CC The invention relates to a multi-epitope polynucleotide (I) encoding
CC cytotoxic T lymphocyte (CTL) and/or helper T lymphocyte (HTL) epitopes.
CC The composition and methods are useful for inducing or stimulating an
CC immune response against hepatitis B virus (HBV), other viruses (e.g.
CC HIV), or against cancers such as prostate cancer or lymphoma. This
CC sequence corresponds to an example of an epigene construct encoded
CC protein used in the invention.

XX
SQ Sequence 456 AA;
SQ

Query Match 15.6%; Score 94; DB 8; Length 456;
Best Local Similarity 85.7%; Pred. No. 0.34;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 40 PAIELPSENERGYIIPHOSSL 60
|
DB 119 PGPGLPSENERGYIIPHOSSL 139

RESULT 11

AAR95607

ID AAR95607 standard; protein; 1085 AA.

XX AAR95607;

XX 12-OCT-1996 (first entry)

XX RRP3 telomerase-associated protein.

XX Yeast; RRP3 protein; telomerase-associated protein; STR7;

XX suppressor of telomeric repression-7; telomerase; ribonucleoprotein;

XX telomere; tumour; pathogen; sperm; ovum; reporter gene; drug screening;

XX antibody; immunosay; antitumour; antiseptic; contraceptive;

XX infertility; diagnostic; gene therapy.

XX Saccharomyces cerevisiae.

XX WO9612811-A2.

XX 02-MAY-1996.

XX 20-OCT-1995; 95WO-US013801.

XX 20-OCT-1994; 94US-00326781.

XX 28-APR-1995; 95US-00431080.

XX (ARCH-) ARCH DEV CORP.

XX Gottschling DE, Singer MS;

XX WPI; 1996-239169/24.

XX N-PSDB; AAT27052.

XX Novel telomerase associated polypeptide(s) and related nucleic acid -

XX useful for detecting e.g. tumour cells or pathogens.

XX Example 10; Page 294-302; 349pp; English.

XX RRP3 telomerase-associated protein (containing fragment STR7 (AAR95606)

XX is found in conjunction with Saccharomyces cerevisiae telomerase, a

XX ribonucleoprotein required for telomere replication. Other proteins

XX associated with telomerase are given in AAR95601-05 (other STR proteins).

XX These proteins combine with telomerase to repress telomere silencing of

XX gene expression. Oligonucleotides from the encoding sequence may be used

XX to detect non-ciliate telomerase-associated genes, e.g. in tumour, may

XX pathogen, sperm or ovum cells. New telomerase-associated sequences may be

XX detected by a reporter gene expression system linked to an expression-

XX representing telomere sequence, and binding compounds, e.g. antibodies, may

XX be detected by complex formation with telomerase components. The products

XX may be used as antitumour, antiseptic or contraceptive agents, in

XX infertility diagnosis, or in gene therapy

XX Sequence 1085 AA;

XX

RESULT 12

ADK62464

ID ADK62464 standard; protein; 1085 AA.

XX ADK62464;

XX 06-MAY-2004 (first entry)

XX Disease treating protein complex-derived protein #358.

XX protein complex; drug target; diagnosis.

XX Undenitified.

XX EP1338608-A2.

XX 27-AUG-2003.

XX 20-DEC-2002; 2002EP-00102902.

XX 20-DEC-2001; 2001EP-00130253.

XX (CELL-) CELLZOME AG.

XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;

XX Marzloch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;

XX Michon A, Leutwein C, Rick J;

XX WPI; 2003-638460/61.

XX N-PSDB; ADK62465.

XX New proteins and protein complexes from eukaryotes, useful as targets in

XX drug screening, or in diagnosing or screening for the presence of a

XX disease or disorder, or a predisposition for developing a disease or

XX disorder in a subject.

XX Disclosure; SEQ ID NO 715; 13pp; English.

XX The invention relates to novel protein complexes comprising a first and a

XX second protein, or its derivative, fragment, homologue or variant. The

XX proteins are selected from given protein complexes, which are not defined

XX in the specification. The variants are encoded by nucleic acids that

XX hybridize to the nucleic acids encoding the proteins under low stringency

XX conditions. The protein complexes are useful as targets for an active

XX agent of a pharmaceutical. These protein complexes are particularly

XX useful as drugs targets for the treatment or preventing of a disease or

XX disorder. The complexes and methods above are useful in diagnosing or

XX screening for the presence of a disease or disorder or a predisposition

XX for developing a disease or disorder in a subject. These are also useful

XX in screening for a drug for treatment or prevention of a disease or

XX disorder. The molecule that modulates the amount, activity or protein

XX components of the complex is useful for the manufacture of a medicament

XX for the treatment or prevention of a disease or disorder. This sequence

XX corresponds to a protein of the invention. (Note: the sequence data for

XX this patent did not form part of the printed specification but was

XX obtained from the EPO in electronic format).

XX Sequence 1085 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

RESULT 13
AAG98701
ID AAG98701 standard; protein; 231 AA.
XX
AC AAG98701;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human cell death protective cDNA clone CNI-00717 ORF5 protein, SEQ:194.
XX
KM Cell death protective; apoptosis; necrosis; human; drug screening;
KM cell death-associated disorder; central nervous system disorder;
KM psychiatric disorder; neurological disorder; ischaemia-related disorder;
KM stroke; cerebral infarction; ischaemic encephalopathy;
KM neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
KM Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
KM vascular disease; ophthalmological disorder; diabetic retinopathy;
KM macular degeneration; hypertension; myocardial infarction;
KM atherosclerosis; respiratory disorder; asthma; transgenic animal;
KM chronic obstructive pulmonary disease; neoplastic condition; cancer;
KM benign tumour; anaemia; gastrointestinal disorder; gastritis;
KM ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KM glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KM Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KM urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.
XX
OS Homo sapiens.
XX
PN WO200145638-A2.
XX
PD 28-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US033547.
XX
PR 14-DEC-1999; 99US-00461697.
XX
PA (COGE-) COGENT NEUROSCIENCE INC.
XX
PI Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;
XX
DR WPI, 2001-390297/41.
XX
DR N-PSDB; AAH84226, AAH84231.
XX
PT Novel protective sequence polynucleotides and polypeptides, used to
PT identify modulators of their expression and activity, which are used in
PT to treat central nervous system conditions, diseases and disorders.
XX
PS Claim 1; Fig 9E; 325pp; English.
XX
CC Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
CC protect against cell death (i.e., apoptosis or necrosis). Sequences
CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
CC while the remaining nucleic acid sequences within the range given above
CC represent the open reading frames (ORFs) of these cDNA clones. Sequences
CC AAG98610-AAG98829 represent the polypeptides encoded by the cell death
CC protective ORFs. The cell death protective cDNA clones are able to
CC prevent, delay or reverse progression through the apoptotic or necrotic
CC pathways when injected into a cell predisposed to or undergoing cell
CC death. The cell death protective nucleic acids and polypeptides can be
CC used in the diagnosis and treatment of disorders associated with cell
CC death, and to screen for compounds which modulate their activity or
CC expression. Such modulators, preferably a small organic molecule, an
CC antibody, a ribozyme, or an antisense molecule, can also be used to treat
CC cell death-related diseases. Such diseases include those associated with
CC the central nervous system including psychiatric or neurological
CC disorders, especially ischaemia-related conditions such as strokes, and
CC also includes neurodegenerative disorders such as Alzheimer's disease,
CC Huntington's disease, or Parkinson's disease. The modulators may also be
CC used to treat infections such as meningitis, malaria, or trypanosomiasis;
CC vascular diseases such as ischaemic encephalopathy or cerebral infarction

CC ; eye conditions such as diabetic retinopathy or macular degeneration;
CC hypertension; myocardial infarction; atherosclerosis; respiratory
CC conditions such as asthma or chronic obstructive pulmonary disease;
CC neoplastic conditions such as cancer or benign tumours; blood cell
CC conditions such as anaemia; gastrointestinal conditions such as gastritis
CC or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney
CC disorders such as glomerulonephritis; cystitis; endometriosis; endocrine
CC disorders such as Grave's disease or Hashimoto's thyroiditis; skin
CC conditions such as dermatitis or urticaria; or immune system disorders
CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may
CC additionally be used to generate animal models of cell death-associated
CC disorders. The present sequence represents a cell death protective
CC polypeptide
XX
SQ Sequence 231 AA;
XX
Query Match 14.8%; Score 89.5; DB 4; Length 231;
Best Local Similarity 29.3%; Pred. No. 0.42;
Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;
QY 3 EQQRLQPKADTKNLRKKKEHGDILADLYGRLEIFALIEPSNENGYIYPHQ--SL 60
DB 71 EVKNEEEDQKEDBEDQNEKEGAG--KEDXDEKEBEDQKDNENGEDEAKKEDGK 127
QY 61 PQDNRGNRSDSKSEISIEKTRRESITTNV--EGRRDIHKHGLEEKDGSIRPEQKE 114
DB 128 GEDGNGNBDGKEKEDKEBEDRKETGVGENEDGKEKDGKKGKGVKVEDKE 183
RESULT 14
AAG98700
ID AAG98700 standard; protein; 232 AA.
XX
AC AAG98700;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human cell death protective cDNA clone CNI-00717 ORF4 protein, SEQ:192.
XX
KM Cell death protective; apoptosis; necrosis; human; drug screening;
KM cell death-associated disorder; central nervous system disorder;
KM psychiatric disorder; neurological disorder; ischaemia-related disorder;
KM stroke; cerebral infarction; ischaemic encephalopathy;
KM neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
KM Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
KM vascular disease; ophthalmological disorder; diabetic retinopathy;
KM macular degeneration; hypertension; myocardial infarction;
KM atherosclerosis; respiratory disorder; asthma; transgenic animal;
KM chronic obstructive pulmonary disease; neoplastic condition; cancer;
KM benign tumour; anaemia; gastrointestinal disorder; gastritis;
KM ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KM glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KM Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KM urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.
XX
OS Homo sapiens.
XX
PN WO200145638-A2.
XX
PD 28-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US033547.
XX
PR 14-DEC-1999; 99US-00461697.
XX
PA (COGE-) COGENT NEUROSCIENCE INC.
XX
PI Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;
XX
DR WPI, 2001-390297/41.
XX
DR N-PSDB; AAH84226, AAH84230.
XX
PT Novel protective sequence polynucleotides and polypeptides, used to

PT identify modulators of their expression and activity, which are used in
PT to treat central nervous system conditions, diseases and disorders.
PS Claim 1; Fig 9D; 325pp; English.
XX Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
CC protect against cell death (i.e., apoptosis or necrosis). Sequences
CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
CC while the remaining nucleic acid sequences within the range given above
CC represent the open reading frames (ORFs) of these cDNA clones. Sequences
CC AAH84132-AAH84370 represent the polypeptides encoded by the cell death
CC protective ORFs. The cell death protective cDNA clones are able to
CC prevent, delay or reverse progression through the apoptotic or necrotic
CC pathways when injected into a cell predisposed to or undergoing cell
CC death. The cell death protective nucleic acids and polypeptides can be
CC used in the diagnosis and treatment of disorders associated with cell
CC death, and to screen for compounds which modulate their activity or
CC expression. Such modulators, preferably a small organic molecule, an
CC antibody, a ribozyme, or an antisense molecule, can also be used to treat
CC cell death-related diseases. Such diseases include those associated with
CC the central nervous system including psychiatric or neurological
CC disorders, especially ischemia-related conditions such as strokes, and
CC also includes neurodegenerative disorders such as Alzheimer's disease,
CC Huntington's disease, or Parkinson's disease. The modulators may also be
CC used to treat infections such as meningitis, malaria, or trypanosomiasis;
CC vascular diseases such as ischemic encephalopathy or cerebral infarction;
CC eye conditions such as diabetic retinopathy or macular degeneration;
CC hypertension; myocardial infarction; atherosclerosis; respiratory
CC conditions such as asthma or chronic obstructive pulmonary disease;
CC neoplastic conditions such as cancers or benign tumours; blood cell
CC conditions such as anaemia; gastrointestinal conditions such as gastritis
CC or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney
CC disorders such as glomerulonephritis; cystitis; endometriosis; skin
CC conditions such as dermatitis or urticaria; or immune system disorders
CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may
CC additionally be used to generate animal models of cell death-associated
CC disorders. The present sequence represents a cell death protective
CC polypeptide
XX
XX Sequence 232 AA;
S0

Query Match 14.8%; Score 89.5; DB 4; Length 232;
Best Local Similarity 29.3%; Pred. No. 0.42;
Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;

QY 3 EQQRDLERKADTKNLERKKEHGDILAEIDYGRLEIPALIEPSENGEYIPIHQ--SSL 60
DB 72 EVKNEEFQKDEDEQNNEKAG---KEDDKENGEDEKDKNENKENGDAKKEDEK 128
QY 61 PODNNGNSRDSKEISIIETKNRESITTVN--EGRRDIHKHLEKKDGSIPEQKE 114
DB 129 GEDGKGNGEDGKEKDEKEDEDRKETGVGKENEDGKKGKDKGVKXDEKE 184

RESULT 15
AAG98699
ID AAG98699 standard; protein; 238 AA.
XX
AC AAG98699;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human cell death protective cDNA clone CNI-00717 ORF3 protein, SEQ.190.
XX
XX Cell death protective; apoptosis; necrosis; human; drug screening;
XX cell death-associated disorder; central nervous system disorder;
XX psychiatric disorder; neurological disorder; ischemia-related disorder;
XX stroke; cerebral infarction; ischemic encephalopathy;
XX neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
XX Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
XX vascular disease; ophthalmological disorder; diabetic retinopathy;

KW macular degeneration; hypertension; myocardial infarction;
KW atherosclerosis; respiratory disorder; asthma; transgenic animal;
KW chronic obstructive pulmonary disease; neoplastic condition; cancer;
KW benign tumour; anaemia; gastrointestinal disorder; gastritis;
KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KW urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.
XX
OS Homo sapiens.
XX
XX WO200145638-A2.
XX
XX 28-JUN-2001.
XX
XX 11-DEC-2000; 2000WO-US033547.
XX
XX 14-DEC-1999; 99US-00461697.
XX
XX (COGE-) COGENT NEUROSCIENCE INC.
XX
XX Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;
XX WPI; 2001-390297/41.
XX N-PSDB; AAH84226, AAH84229.
XX
XX Novel protective sequence polynucleotides and polypeptides, used to
PT identify modulators of their expression and activity, which are used in
PT to treat central nervous system conditions, diseases and disorders.
XX
XX Claim 1; Fig 9C; 325pp; English.
XX
XX Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
CC protect against cell death (i.e., apoptosis or necrosis). Sequences
CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
CC while the remaining nucleic acid sequences within the range given above
CC represent the open reading frames (ORFs) of these cDNA clones. Sequences
CC AAH84132-AAH84370 represent the polypeptides encoded by the cell death
CC protective ORFs. The cell death protective cDNA clones are able to
CC prevent, delay or reverse progression through the apoptotic or necrotic
CC pathways when injected into a cell predisposed to or undergoing cell
CC death. The cell death protective nucleic acids and polypeptides can be
CC used in the diagnosis and treatment of disorders associated with cell
CC death, and to screen for compounds which modulate their activity or
CC expression. Such modulators, preferably a small organic molecule, an
CC antibody, a ribozyme, or an antisense molecule, can also be used to treat
CC cell death-related diseases. Such diseases include those associated with
CC the central nervous system including psychiatric or neurological
CC disorders, especially ischemia-related conditions such as strokes, and
CC also includes neurodegenerative disorders such as Alzheimer's disease,
CC Huntington's disease, or Parkinson's disease. The modulators may also be
CC used to treat infections such as meningitis, malaria, or trypanosomiasis;
CC vascular diseases such as ischemic encephalopathy or cerebral infarction;
CC eye conditions such as diabetic retinopathy or macular degeneration;
CC hypertension; myocardial infarction; atherosclerosis; respiratory
CC conditions such as asthma or chronic obstructive pulmonary disease;
CC neoplastic conditions such as cancers or benign tumours; blood cell
CC conditions such as anaemia; gastrointestinal conditions such as gastritis
CC or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney
CC disorders such as glomerulonephritis; cystitis; endometriosis; endocrine
CC disorders such as Grave's disease or Hashimoto's thyroiditis; skin
CC conditions such as dermatitis or urticaria; or immune system disorders
CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may
CC additionally be used to generate animal models of cell death-associated
CC disorders. The present sequence represents a cell death protective
CC polypeptide
XX
XX Sequence 238 AA;
S0

Query Match 14.8%; Score 89.5; DB 4; Length 238;
Best Local Similarity 29.3%; Pred. No. 0.44;
Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:16:04 ; Search time 11.6189 Seconds
(without alignments)
832.525 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LOEQQRDLQKRAADTKKMLE.....GHLEKKKDSIKPEKKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/prodataa/1/iaa/5.COMB.pep.*
2: /cgn2_6/prodataa/1/iaa/6.COMB.pep.*
3: /cgn2_6/prodataa/1/iaa/H.COMB.pep.*
4: /cgn2_6/prodataa/1/iaa/PCTUS.COMB.pep.*
5: /cgn2_6/prodataa/1/iaa/RE.COMB.pep.*
6: /cgn2_6/prodataa/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	603	100.0	117	2	US-08-098-327E-20	Sequence 20, Appl
2	603	100.0	117	2	US-08-462-625-20	Sequence 20, Appl
3	603	100.0	117	2	US-08-098-327E-31	Sequence 31, Appl
4	603	100.0	117	2	US-08-462-625-31	Sequence 31, Appl
5	554	91.9	107	2	US-08-098-327E-19	Sequence 19, Appl
6	554	91.9	107	2	US-08-462-625-19	Sequence 19, Appl
7	154	25.5	31	2	US-08-098-327E-23	Sequence 23, Appl
8	154	25.5	31	2	US-08-462-625-23	Sequence 23, Appl
9	141	23.4	27	2	US-08-098-327E-21	Sequence 21, Appl
10	141	23.4	27	2	US-08-462-625-21	Sequence 21, Appl
11	97	16.1	709	2	US-09-248-796A-19045	Sequence 19045, A
12	90.5	15.0	1085	1	US-08-431-080-28	Sequence 28, Appl
13	90.5	15.0	1085	1	US-08-938-534-28	Sequence 28, Appl
14	90.5	15.0	1085	2	US-09-345-294-28	Sequence 28, Appl
15	89.5	14.8	231	2	US-09-461-697-194	Sequence 194, App
16	89.5	14.8	231	2	US-09-461-697-192	Sequence 192, App
17	89.5	14.8	238	2	US-09-461-697-190	Sequence 190, App
18	89.5	14.8	257	2	US-09-461-697-188	Sequence 188, App
19	89.5	14.8	272	2	US-09-461-697-186	Sequence 186, App
20	86	14.3	447	2	US-09-710-279-2900	Sequence 2900, Ap
21	86	14.3	450	2	US-09-710-279-3226	Sequence 3226, Ap
22	85.5	14.2	765	2	US-09-538-092-906	Sequence 906, App
23	85.5	14.2	765	2	US-09-882-274-2	Sequence 2, Appl
24	85	14.1	15	2	US-09-009-953-74	Sequence 74, Appl
25	84.5	14.0	150	2	US-09-395-689-2	Sequence 2, Appl
26	84.5	14.0	300	2	US-09-395-689-1	Sequence 1, Appl
27	84.5	14.0	584	2	US-09-949-016-8587	Sequence 8587, Ap

28	84.5	14.0	585	2	US-09-949-016-6627	Sequence 6627, Ap
29	84.5	14.0	765	1	US-08-663-112-2	Sequence 2, Appl
30	84	13.9	24	2	US-08-098-327E-22	Sequence 22, Appl
31	84	13.9	24	2	US-08-462-625-22	Sequence 22, Appl
32	84	13.9	1761	2	US-09-561-709B-1	Sequence 1, Appl
33	83.5	13.8	1025	2	US-09-538-092-334	Sequence 334, App
34	82.5	13.7	1027	2	US-09-514-259-27	Sequence 27, Appl
35	80.5	13.3	129	2	US-09-513-999C-5707	Sequence 5707, Ap
36	80.5	13.3	346	2	US-09-902-540-12787	Sequence 12787, A
37	80.5	13.3	1032	2	US-09-914-255-26	Sequence 26, Appl
38	80.5	13.3	1032	2	US-09-538-092-1293	Sequence 1293, Ap
39	79	13.1	140	2	US-09-710-279-2268	Sequence 2268, Ap
40	79	13.1	140	2	US-09-710-279-2438	Sequence 2438, Ap
41	79	13.1	147	2	US-09-134-001C-5518	Sequence 5518, Ap
42	79	13.1	538	2	US-09-252-991A-23551	Sequence 23551, A
43	79	13.1	708	2	US-08-235-835C-76	Sequence 76, Appl
44	78	12.9	452	2	US-09-248-796A-14553	Sequence 14553, A
45	78	12.9	1270	2	US-09-538-092-1321	Sequence 1321, Ap

ALIGNMENTS

RESULT 1
US-08-098-327E-20
; Sequence 20, Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUIHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SRO ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-098-327E-20
Query Match 100.0%; Score 603; DB 2; Length 117;

Best Local Similarity 100.0%; Pred. No. 2,4e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
DB 1 LOEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
QY 61 PODNRGNSRDSKESITIEKTNRESITTNVEGRDIIHGHLBEKKDGSIKPEOKEDKS 117
DB 61 PODNRGNSRDSKESITIEKTNRESITTNVEGRDIIHGHLBEKKDGSIKPEOKEDKS 117

RESULT 2

US-08-462-625-20
; Sequence 20, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,625
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/098,327
; FILING DATE: 24-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-462-625-20

Query Match 100.0%; Score 603; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 2,4e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
DB 1 LOEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
QY 61 PODNRGNSRDSKESITIEKTNRESITTNVEGRDIIHGHLBEKKDGSIKPEOKEDKS 117

DB 61 PODNRGNSRDSKESITIEKTNRESITTNVEGRDIIHGHLBEKKDGSIKPEOKEDKS 117

RESULT 3

US-08-098-327E-31
; Sequence 31, Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; US-08-098-327E-31

Query Match 100.0%; Score 603; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 9,4e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 60
DB 200 LOEQORDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPALIEPSENERGYIIPHOSSL 259
QY 61 PODNRGNSRDSKESITIEKTNRESITTNVEGRDIIHGHLBEKKDGSIKPEOKEDKS 117
DB 260 PODNRGNSRDSKESITIEKTNRESITTNVEGRDIIHGHLBEKKDGSIKPEOKEDKS 316

RESULT 4

US-08-462-625-31
; Sequence 31, Application US/08462625
; Patent No. 6319502
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; APPLICANT: DRUILHE, Pierre

TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-31
Query Match 100.0%; Score 603; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 9.4e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQEQRDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPIALPSENERGYIIPHOSSL 60
DB 200 LQEQRDLEQRKADTKNLERKKEHGDILAEDLYGRLEIPIALPSENERGYIIPHOSSL 259
QY 61 PQQNRGNSRDSKSIISTEKTNRSSITTVNGRRDINHGHLEKKDGSIRPEOKEDKS 117
DB 260 PQQNRGNSRDSKSIISTEKTNRSSITTVNGRRDINHGHLEKKDGSIRPEOKEDKS 316
RESULT 5
US-08-098-327E-19
Sequence 19, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria

STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-19
Query Match 91.9%; Score 554; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 8.4e-56;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RKADTKNLERKKEHGDILAEDLYGRLEIPIALPSENERGYIIPHOSSLPQNRNSRD 70
DB 1 RKADTKNLERKKEHGDILAEDLYGRLEIPIALPSENERGYIIPHOSSLPQNRNSRD 60
QY 71 SKSISTEKTNRSSITTVNGRRDINHGHLEKKDGSIRPEOKEDKS 117
DB 61 SKSISTEKTNRSSITTVNGRRDINHGHLEKKDGSIRPEOKEDKS 107
RESULT 6
US-08-462-625-19
Sequence 19, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-19

Query Match 91.9%; Score 554; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 8,4e-56;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADPKKLEKKEKGGDILAEEDLYGRLEIPALPEPSENERGYIIPQSSLPQDRNGNSRD 70
Db 1 RKADPKKLEKKEKGGDILAEEDLYGRLEIPALPEPSENERGYIIPQSSLPQDRNGNSRD 60

Qy 71 SKEISIIKTNRESITTVNVEGRDIIKHGHLKLEKKDGSIKPEQKEDKS 117
Db 61 SKEISIIKTNRESITTVNVEGRDIIKHGHLKLEKKDGSIKPEQKEDKS 107

RESULT 7
US-08-098-327E-23
Sequence 23, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-23

Query Match 25.5%; Score 154; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 9e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 NSRDSKESIIKTNRESITTVNVEGRDIIKH 97
Db 1 NSRDSKESIIKTNRESITTVNVEGRDIIKH 31

RESULT 8
US-08-462-625-23
Sequence 23, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992
US-08-462-625-23

Query Match 25.5%; Score 154; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 9e-11;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 NSRDSKEISIIIEKTNRESITTTNVEGRDHIK 97
DB 1 NSRDSKEISIIIEKTNRESITTTNVEGRDHIK 31

RESULT 9
US-08-098-327E-21
Sequence 21, Application US/08098327E

PATENT INFORMATION:
PATENT NO. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: DRUITH, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-21

Query Match 23.4%; Score 141; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DTKKNLEKKKHGDIABDIYGRLEIP 40
DB 1 DTKKNLEKKKHGDIABDIYGRLEIP 27

RESULT 10
US-08-462-625-21
Sequence 21, Application US/08462625

PATENT NO. 6319502
GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
TITLE OF INVENTION: DRUITH, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-21

Query Match 23.4%; Score 141; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DTKKNLEKKKHGDIABDIYGRLEIP 40
DB 1 DTKKNLEKKKHGDIABDIYGRLEIP 27

RESULT 11
US-09-248-796A-19045
Sequence 19045, Application US/09248796A

PATENT NO. 6747137
GENERAL INFORMATION:
APPLICANT: Kelch Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196,132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 19045
LENGTH: 709
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-19045

Query Match
Best Local Similarity 24.7%; Pred. No. 0.023;
Matches 37; Conservative 29; Mismatches 36; Indels 48; Gaps 9;

QY 11 RKADTKKLEKKKEKHGDI-----LAEDLYGRLEIPA-----IELPSENERGYI 54
Db 462 KEVDTDLSIQEKELFQIKPEVNDKDEKQIKEDLV-KVETPSKDEGPAKVEVPSKDETPRKI 520
QY 55 PHQSSLPQDNNGNSRDSK-----EISIEKTNREST-----TTNVEGRD 94
Db 521 QNIDS-TTDN-GKAKDEKDNVSKVDSNVDTKVDQDQGIKVEKEPEKKDDDEGRKD 578
QY 95 -----IHKGHLEKKKDGSIKPE---OKED 115
Db 579 NEKVTVHKNEEDKDKGVNEAKDKVKQKD 608

RESULT 12
US-08-431-080-28
Sequence 28, Application US/08431080
Patent No. 3698686
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431.080
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-431-080-28

Query Match
Best Local Similarity 15.0%; Score 90.5; DB 1; Length 1085;
Matches 31; Conservative 21; Mismatches 54; Indels 43; Gaps 4;

QY 3 EQQRDLQRKADTKKLEKKKE-----HGD 27
Db 865 DKMKELQRKKTTEKKRQLKKKKLLKIRKQKAIKEQETMNLQINGHEIIGNNNSHD 924
QY 28 I-LAEDLYGRLEIPAILEPSE-NEGYIIPHQSSLPQDNNGNSRDSKEI-----SIEK 79
Db 925 INTGDTFTTNENTPMNELPSHAPEDASLIPIHNSDLAVSNTRKSTKSVGLDEIHEILGK 984
QY 80 TNRSEITTVNTEGRDIIKHGLEKKDGSI 108
Db 985 DENDLSVGDINGIDAGQGHVIEDTDADI 1013

RESULT 13
US-08-938-534-28
Sequence 28, Application US/08938534
Patent No. 5916752
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938.534
FILING DATE: 26-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,080
FILING DATE:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-938-534-28

Query Match
Best Local Similarity 15.0%; Score 90.5; DB 1; Length 1085;
Matches 31; Conservative 21; Mismatches 54; Indels 43; Gaps 4;

QY 3 EQQRDLQRKADTKKLEKKKE-----HGD 27
Db 865 DKMKELQRKKTTEKKRQLKKKKLLKIRKQKAIKEQETMNLQINGHEIIGNNNSHD 924
QY 28 I-LAEDLYGRLEIPAILEPSE-NEGYIIPHQSSLPQDNNGNSRDSKEI-----SIEK 79
Db 925 INTGDTFTTNENTPMNELPSHAPEDASLIPIHNSDLAVSNTRKSTKSVGLDEIHEILGK 984
QY 80 TNRSEITTVNTEGRDIIKHGLEKKDGSI 108

Db 985 DENDLLSVGDINGYDAQGHVIEDTDADI 1013

RESULT 14

US-09-345-294-28
Sequence 28, Application US/09345294
Patent No. 6387619

GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.

Singer, Miriam S.

TITLE OF INVENTION: Telomerase Compositions and Methods

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TEXAS

COUNTRY: UNITED STATES OF AMERICA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/345,294

FILING DATE: 30-Jun-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/431,080

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.

REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: ARCD:155/PAR

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 1085 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 28:

US-09-345-294-28

Query Match 15.0%; Score 90.5; DB 2; Length 1085;

Best Local Similarity 20.8%; Pred. No. 0.23;

Matches 31; Conservative 21; Mismatches 54; Indels 43; Gaps 4;

QY 3 EQQRDLQQRKADTKKLEKKKE-----HGD 27

Db 865 DKKEELQRRKTEKKRQOLKKKKLTKRKQKAKKEQETMNLQLGNGHIIIGNNNSHSD 924

QY 28 I-LAEDLYGRLEIPIALEPSE-NEGYIIPHOSSLPODNRGNSRDSKEI-----SIIEX 79

Db 925 INTGDTPTTMENTPMNLPSHAPEDASLIHNSDLAVDSNTRKNSSTKSVGLDEIHEILGK 984

QY 80 TNRESITTVNGRRDIHKHLEKKKGGSI 108

Db 985 DENDLLSVGDINGYDAQGHVIEDTDADI 1013

RESULT 15

US-09-461-697-194

Sequence 194, Application US/09461697

Patent No. 6277974

GENERAL INFORMATION:

APPLICANT: COGENT NEUROSCIENCE, Inc.

APPLICANT: LO, Donald C.

APPLICANT: Barney, Shawn

APPLICANT: Thomas, Mary Beth

APPLICANT: Portbury, Stuart D.

APPLICANT: Putnam, Kaeturi

APPLICANT: Katz, Lawrence C.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING

TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING

TITLE OF INVENTION: CELL DEATH

FILE REFERENCE: 10001-005-999

CURRENT APPLICATION NUMBER: US/09/461,697

CURRENT FILING DATE: 1999-12-14

NUMBER OF SEQ ID NOS: 466

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 194

LENGTH: 231

TYPE: PRT

ORGANISM: Homo sapiens

US-09-461-697-194

Query Match 14.8%; Score 89.5; DB 2; Length 231;

Best Local Similarity 29.3%; Pred. No. 0.036;

Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;

QY 3 EQQRDLQQRKADTKKLEKKKEHGDILAEPLPALEPSENEGYIIPHO--SSL 60

Db 71 EYVNEEDQKEDBEDQNEKGEAG--KEDKXKGEDEGDKNGNEKGEADAKKEDKK 127

QY 61 PODNRGNSRDSKEISIIETKTNRBSITTV--EGRRDIHKHLEKKKGGSIKPEQKE 114

Db 128 GEDGKNGEDGKEKGEDEKEDRKETGVGKENEDGKGGKGGDVYKDEKE 183

Search completed: December 7, 2005, 07:22:11
Job time: 12.6189 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2005, 07:18:34 ; Search time 30.5335 Seconds
(without alignments)
1601.063 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LDEQQRDLQRKADTKKNLE.....GHLEKKDSIKPEQKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603	100.0	117	3	US-09-837-344-20
2	603	100.0	117	3	US-09-900-963-20
3	603	100.0	117	3	US-09-837-344-31
4	603	100.0	117	3	US-09-837-344-31
5	602	99.8	1909	5	US-10-732-923-3341
6	554	91.9	107	3	US-09-837-344-19
7	554	91.9	107	3	US-09-900-963-19
8	154	25.5	31	3	US-09-837-344-23
9	154	25.5	31	3	US-09-900-963-23
10	141	23.4	27	3	US-09-837-344-21
11	141	23.4	27	3	US-09-900-963-21
12	94	15.6	456	3	US-09-894-018-121
13	94	15.6	456	3	US-10-474-960A-121
14	90	14.9	756	5	US-10-756-149-5156
15	89.5	14.8	231	3	US-09-922-261-194
16	89.5	14.8	231	3	US-09-922-261-192
17	89.5	14.8	238	3	US-09-922-261-190
18	89.5	14.8	257	3	US-09-922-261-188
19	89.5	14.8	272	3	US-09-922-261-186
20	87	14.4	550	4	US-10-369-493-5995
21	85.5	14.2	765	4	US-09-882-274-2
22	85.5	14.2	765	4	US-10-408-765A-1149
23	85.5	14.2	765	4	US-10-408-577-679
24	85	14.1	15	4	US-10-103-395-74
25	84.5	14.0	242	5	US-10-450-763-32854
26	84.5	14.0	585	5	US-10-631-467-870
27	84	13.9	24	3	US-09-837-344-22

28	84	13.9	24	3	US-09-900-963-22	Sequence 22, Appl
29	83.5	13.8	257	4	US-10-425-114-58860	Sequence 58860, A
30	83.5	13.8	399	4	US-10-425-114-69744	Sequence 69744, A
31	83.5	13.8	399	4	US-10-425-115-226794	Sequence 226794, A
32	83.5	13.8	419	4	US-10-425-115-226797	Sequence 226797, A
33	83.5	13.8	807	5	US-10-511-698-39	Sequence 39, Appl
34	83.5	13.8	903	4	US-10-282-122A-52328	Sequence 52328, A
35	83	13.8	2683	4	US-10-437-963-185370	Sequence 185370, A
36	82.5	13.7	1027	4	US-10-080-608A-27	Sequence 27, Appl
37	82.5	13.7	1027	4	US-10-370-685-116	Sequence 116, Appl
38	82.5	13.7	1264	4	US-10-369-493-3742	Sequence 3742, Ap
39	82.5	13.7	1642	5	US-10-732-923-8639	Sequence 8639, Ap
40	82	13.6	472	3	US-09-953-407-2	Sequence 2, Appl1
41	82	13.6	1429	5	US-10-450-763-35896	Sequence 35896, A
42	81.5	13.5	485	3	US-09-769-736-72	Sequence 72, Appl
43	81.5	13.5	980	5	US-10-732-923-3343	Sequence 3343, Ap
44	81.5	13.5	2447	5	US-10-450-763-52739	Sequence 52739, A
45	81.5	13.5	2447	5	US-10-450-763-52739	Sequence 52739, A

ALIGNMENTS

RESULT 1
US-09-837-344-20
; Sequence 20, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-09-837-344-20

Query Match 100.0%; Score 603; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 6.8e-51;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 60
1 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 60
Db 1 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 60
61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117
61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117
Db 61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117

RESULT 2

US-09-900-963-20
Sequence 20, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSER: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22133-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/900,963

FILING DATE: 10-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/098,327

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-045

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 117 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-09-900-963-20

Query Match 100.0%; Score 603; DB 3; Length 117;
Best Local Similarity 100.0%; Pred. No. 6.8e-51;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 60
1 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 60
Db 1 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 60
61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117
61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117
Db 61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117

Qy 61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117
61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117
Db 61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117

RESULT 3

US-09-837-344-31
Sequence 31, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSER: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22133-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/837,344

FILING DATE: 19-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/462,625

FILING DATE: <Unknown>

APPLICATION NUMBER: FR 91 01286

FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 316 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

PUBLICATION INFORMATION:

DOCUMENT NUMBER: WO 92/13884

PUBLICATION DATE: 20-AUG-1992

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-837-344-31

Query Match 100.0%; Score 603; DB 3; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.4e-50;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 60
1 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 60
Db 200 LOEQRDLQKADTKKLNLEKKEHGDILAEGLYGRLEIPALPSPSENERGYIIPHOSSL 259
61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117
61 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 117
Db 260 PQRNRGSRDSKESIIIEKTNRESITTVNVEGRDIIHGHLLEKKDGSIKPEQKEDKS 316

RESULT 4

US-09-900-963-31
Sequence 31, Application US/09900963

Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: HEPATIC SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-900-963-31
Query Match 100.0%; Score 603; DB 3; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.4e-50;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAE DLYGRLEIPALIELPSENERGYIIPHOSSL 60
DB 200 LOEQORDLEQRKADTKNLERKKEHGDILAE DLYGRLEIPALIELPSENERGYIIPHOSSL 259
QY 61 PDNRGNSRDSKEISIIETKNTRESITTVNVEGRDINHKGHLEKKDGSIKPEQKEDKS 117
DB 260 PDNRGNSRDSKEISIIETKNTRESITTVNVEGRDINHKGHLEKKDGSIKPEQKEDKS 316
RESULT 5
US-10-732-923-3341
Sequence 3341, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149

SEQ ID NO 3341
LENGTH: 1909
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-732-923-3341
Query Match 99.8%; Score 602; DB 5; Length 1909;
Best Local Similarity 99.1%; Pred. No. 2.9e-49;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LOEQORDLEQRKADTKNLERKKEHGDILAE DLYGRLEIPALIELPSENERGYIIPHOSSL 60
DB 1620 LOEQORDLEQRKADTKNLERKKEHGDILAE DLYGRLEIPALIELPSENERGYIIPHOSSL 1679
QY 61 PDNRGNSRDSKEISIIETKNTRESITTVNVEGRDINHKGHLEKKDGSIKPEQKEDKS 117
DB 1680 PDNRGNSRDSKEISIIETKNTRESITTVNVEGRDINHKGHLEKKDGSIKPEQKEDKS 1736
RESULT 6
US-09-837-344-19
Sequence 19, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: HEPATIC SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-837-344-19
Query Match 91.9%; Score 554; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.6e-46;

Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEGGDILADLYGRLEIPALELPSENERGYIIPHOSLIPODNRGNSRD 70
Db 1 RKADTKNLERKKEGGDILADLYGRLEIPALELPSENERGYIIPHOSLIPODNRGNSRD 60
Qy 71 SKEISIIKTNRESITTVNVEGRDIHKHGLEEKDGSIKPKOKEDKS 117
Db 61 SKEISIIKTNRESITTVNVEGRDIHKHGLEEKDGSIKPKOKEDKS 107

RESULT 7

US-09-900-963-19
; Sequence 19, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,963
; FILING DATE: 10-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/098,327
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-900-963-19

Query Match 91.9%; Score 554; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.6e-46;

Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RKADTKNLERKKEGGDILADLYGRLEIPALELPSENERGYIIPHOSLIPODNRGNSRD 70
Db 1 RKADTKNLERKKEGGDILADLYGRLEIPALELPSENERGYIIPHOSLIPODNRGNSRD 60
Qy 71 SKEISIIKTNRESITTVNVEGRDIHKHGLEEKDGSIKPKOKEDKS 117
Db 61 SKEISIIKTNRESITTVNVEGRDIHKHGLEEKDGSIKPKOKEDKS 107

RESULT 8
US-09-837-344-23
; Sequence 23, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: WO 92/13884
; PUBLICATION DATE: 20-AUG-1992
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-837-344-23

Query Match 25.5%; Score 154; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 6.6e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 NSRDSKESIIKTNRESITTVNVEGRDIHK 97
Db 1 NSRDSKESIIKTNRESITTVNVEGRDIHK 31

RESULT 9

US-09-900-963-23
; Sequence 23, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUILHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-900-963-23
Query Match 25.5%; Score 154; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 6,6e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 67 NSRDSKEISIIETKNTRESITTVNVEGRDIIK 97
DB 1 NSRDSKEISIIETKNTRESITTVNVEGRDIIK 31
RESULT 10
US-09-837-344-21
Sequence 21, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344

FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-837-344-21
Query Match 23.4%; Score 141; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 14 DTKNLERKKEHGDIIAEDLYGRLEIP 40
DB 1 DTKNLERKKEHGDIIAEDLYGRLEIP 27
RESULT 11
US-09-900-963-21
Sequence 21, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 27 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; PUBLICATION INFORMATION:
 ; DOCUMENT NUMBER: WO 92/13884
 ; PUBLICATION DATE: 20-AUG-1992
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
 US-09-900-963-21

Query Match 23.4%; Score 141; DB 3; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DTKKNLERKKEHGDILADLYGRLEIP 40
 Db 1 DTKKNLERKKEHGDILADLYGRLEIP 27

RESULT 12
 US-09-894-018-121
 ; Sequence 121, Application US/09894018
 ; Patent No. US20020119127A1

; GENERAL INFORMATION:
 ; APPLICANT: EPIMUNE, Inc.
 ; APPLICANT: Sette, Alessandro
 ; APPLICANT: Chestnut, Robert
 ; APPLICANT: Livingston, Brian
 ; APPLICANT: Baker, Deniaw
 ; APPLICANT: Newman, Mark
 ; APPLICANT: Brown, David
 ; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
 ; FILE REFERENCE: 39963-2003.00
 ; CURRENT APPLICATION NUMBER: US/09/894,018
 ; CURRENT FILING DATE: 2001-06-27
 ; PRIOR APPLICATION NUMBER: PCT/US00/35568
 ; PRIOR FILING DATE: 2000-12-28
 ; PRIOR APPLICATION NUMBER: US 60/173,390
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: US 60/284,221
 ; PRIOR FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 368
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 121
 ; LENGTH: 456
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: PFCITL/HTL(N)
 US-09-894-018-121

Query Match 15.6%; Score 94; DB 3; Length 456;
 Best Local Similarity 85.7%; Pred. No. 1.3;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 40 PAIELPSENERGYIIPHOSSL 60
 Db 119 PGPGLPSENERGYIIPHOSSL 139

RESULT 13
 US-10-474-960A-121
 ; Sequence 121, Application US/10474960A
 ; Publication No. US20040248113A1

; GENERAL INFORMATION:
 ; APPLICANT: Sette, Alessandro
 ; APPLICANT: Chestnut, Robert
 ; APPLICANT: Livingston, Brian
 ; APPLICANT: Baker, Deniaw

; APPLICANT: Newman, Mark
 ; APPLICANT: Brown, David
 ; TITLE OF INVENTION: Method and System for Optimizing Multi-epitope Nucleic
 ; TITLE OF INVENTION: Acid Constructs and Peptides Encoded Thereby
 ; FILE REFERENCE: 2060.0320004
 ; CURRENT APPLICATION NUMBER: US/10/474,960A
 ; CURRENT FILING DATE: 2003-10-16
 ; PRIOR APPLICATION NUMBER: PCT/US02/09877
 ; PRIOR FILING DATE: 2002-03-28
 ; PRIOR APPLICATION NUMBER: US 09/894,018
 ; PRIOR FILING DATE: 2001-06-27
 ; PRIOR APPLICATION NUMBER: US 60/284,221
 ; PRIOR FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 419
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 121
 ; LENGTH: 456
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: PFCITL/HTL(N)
 US-10-474-960A-121

Query Match 15.6%; Score 94; DB 5; Length 456;
 Best Local Similarity 85.7%; Pred. No. 1.3;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 40 PAIELPSENERGYIIPHOSSL 60
 Db 119 PGPGLPSENERGYIIPHOSSL 139

RESULT 14
 US-10-756-149-5156
 ; Sequence 5156, Application US/10756149
 ; Publication No. US20050181375A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Zlotnik, Albert
 ; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
 ; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
 ; FILE REFERENCE: file
 ; CURRENT APPLICATION NUMBER: US/10/756,149
 ; CURRENT FILING DATE: 2004-01-12
 ; NUMBER OF SEQ ID NOS: 5818
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 5156
 ; LENGTH: 756
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-10-756-149-5156

Query Match 14.9%; Score 90; DB 5; Length 756;
 Best Local Similarity 23.5%; Pred. No. 6.2;
 Matches 31; Conservative 24; Mismatches 53; Indels 24; Gaps 4;

QY 2 QEQGRDLQKADKKNLERKKEHGDILADLYGRLE-IPAIELPSENERG----- 51
 Db 594 QENRRMGDRLSHERHOFQDKKATQELIEDLRKQLEHLQULKLEAEGRGRSSMGLOE 653

QY 52 YIIPHOSS-----LPQNRGNSRDSKEISIEKTNRESITTNVEGRDIIHKHGLEEK 103
 Db 654 YHSRARESELEQEVRLKQDRNLKEQNEEL-----NGQITLTSIQAKSLFSTAFSES 707

QY 104 KQGSIKPEQKED 115
 Db 708 LAEISSYSRDE 719

RESULT 15
 US-09-922-261-194
 ; Sequence 194, Application US/09922261
 ; Patent No. US20020111471A1

```

; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Piranam, Kaesturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-261-194

```

```

Query Match      14.8%; Score 89.5; DB 3; Length 231;
Best Local Similarity 29.3%; Pred. No. 1.6;
Matches 34; Conservative 15; Mismatches 60; Indels 7; Gaps 3;

QY      3  EGOQDLQKADTKKTLERKKEHGDILAEDLYGRLEIPALTELPSENERGYIYPHQ--SSL 60
      71  EVKNEEDQKEDKEDQNEKEGANG--KEDKDEKGEEDGKEDKNGNEKGEADAKEDGKK 127

QY      61  PODNRGNSRDSKESISIEKTNRESITTNV--EGRRDIHGKHLSEKKDGSIKPEQKE 114
      128  GEDGKNGEDGKEKGEDEKEDRKETGVGKENEDGKKGKKGKDVVKYKEDDEKE 183
DB

```

Search completed: December 7, 2005, 07:30:34
 Job time : 31.5335 sec

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 7, 2005, 07:21:30 ; Search time 2.70208 Seconds
(without alignments)
241.806 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LOEQORDLEQRKADTKKNE.....GHLEKKGGSIXPEKEDKS 117

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_New.*
1: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB pep.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB pep.*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB pep.*
4: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB pep.*
5: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB pep.*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB pep.*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB pep.*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	14.3	447	US-10-793-626-2900	Sequence 2900, Ap
2	86	14.3	450	US-10-793-626-3226	Sequence 3226, Ap
3	79.5	13.2	2432	US-10-821-234-899	Sequence 899, App
4	79	13.1	140	US-10-793-626-2268	Sequence 2268, Ap
5	79	13.1	140	US-10-793-626-2438	Sequence 2438, Ap
6	79	13.1	795	US-10-770-726-49	Sequence 49, App1
7	76.5	12.7	677	US-10-982-545-12	Sequence 12, App1
8	75	12.4	485	US-10-508-263-22	Sequence 22, App1
9	72.5	12.0	491	US-10-793-626-2808	Sequence 2808, Ap
10	71.5	11.9	212	US-10-467-657-5486	Sequence 5486, Ap
11	71.5	11.9	212	US-10-467-657-7404	Sequence 7404, Ap
12	71	11.8	328	US-10-467-657-7710	Sequence 7710, Ap
13	69	11.4	793	US-11-060-914-2	Sequence 2, App11
14	68	11.3	531	US-11-060-914-4	Sequence 4, App11
15	67	11.1	369	US-10-467-657-1336	Sequence 1336, Ap
16	67	11.1	904	US-10-507-275-5	Sequence 5, App11
17	66	10.9	495	US-10-467-657-3844	Sequence 3844, Ap
18	66	10.9	658	US-10-467-657-4782	Sequence 4782, Ap
19	66	10.9	676	US-11-135-835-28	Sequence 28, App1
20	66	10.9	717	US-11-135-835-29	Sequence 29, App1
21	66	10.9	862	US-11-077-550-171	Sequence 171, App
22	66	10.9	862	US-11-077-550-173	Sequence 173, App
23	65.5	10.9	257	US-10-793-626-126	Sequence 126, App
24	65.5	10.9	397	US-10-793-626-1244	Sequence 1244, Ap
25	65.5	10.9	578	US-10-821-234-1039	Sequence 1039, Ap

26	65	10.8	360	US-10-878-556A-134	Sequence 134, App
27	65	10.8	360	US-11-186-284-115	Sequence 115, App
28	65	10.8	790	US-10-763-712A-75	Sequence 75, App1
29	64.5	10.7	684	US-10-793-626-2098	Sequence 2098, Ap
30	64.5	10.7	1142	US-11-109-156-22	Sequence 22, App1
31	64	10.6	904	US-11-087-227-12	Sequence 12, App1
32	63.5	10.5	215	US-10-793-626-1650	Sequence 1650, Ap
33	63.5	10.5	286	US-10-793-626-2192	Sequence 2192, Ap
34	63.5	10.5	612	US-10-467-657-3988	Sequence 3988, Ap
35	63.5	10.5	664	US-10-878-556A-44	Sequence 44, App1
36	63	10.4	963	US-10-467-962B-2	Sequence 2, App11
37	63	10.4	1404	US-10-878-556A-169	Sequence 169, App
38	62.5	10.4	505	US-10-821-234-1287	Sequence 1287, Ap
39	62	10.3	548	US-11-077-550-24	Sequence 24, App1
40	62	10.3	858	US-11-077-550-22	Sequence 22, App1
41	62	10.3	1169	US-11-077-550-20	Sequence 20, App1
42	61.5	10.2	387	US-10-485-517-132	Sequence 132, App
43	61.5	10.2	451	US-10-467-657-366	Sequence 366, App
44	61.5	10.2	635	US-10-821-234-1573	Sequence 1573, Ap
45	61.5	10.2	858	US-10-645-441-25	Sequence 25, App1

ALIGNMENTS

RESULT 1
US-10-793-626-2900
; Sequence 2900, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2900
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2900

Query Match 14.3%; Score 86; DB 6; Length 447;
Best Local Similarity 23.8%; Pred. NO. 0.18; Mismatches 26; Gaps 4;
Matches 29; Conservative 24; Mismatches 43; Indels 26; Gaps 4;

QY 1 LOEQORDLEQRKADTKKNE-----RKKEGDIADLYGRLETPALEPSENERGYYP 55
DB 15 ISKAQAVAEKGDLETFANLKADIDAKKEYEL--EQLSKEIFASA-----P 59
QY 56 HGSLLPDQNRGNSRDSKSEIIEIKTNRESITTVNBSGRDIIHGKHEKKDGSIKPEOKED 115
DB 60 KQDEPPKDGAEVEDKNGSGSESENKPSDDEPEGTSD-----BEKPDADKPDKPE 113

QY 116 KS 117
DB 114 ET 115

RESULT 2
US-10-793-626-3226
; Sequence 3226, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US

```
;; CURRENT APPLICATION NUMBER: US/10/793,626
;; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;; FILE REFERENCE: PU3480US
;; CURRENT FILING DATE: 2004-03-04
;; PRIOR APPLICATION NUMBER: 60/164,258
;; PRIOR FILING DATE: 1999-11-09
;; NUMBER OF SEQ ID NOS: 4472
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 3226
;; LENGTH: 450
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: synthetic
;; OTHER INFORMATION: amino acid sequence
US-10-793-626-3226

Query Match          14.3%; Score 86; DB 6; Length 450;
Best Local Similarity 23.8%; Pred. No. 0.18;
Matches 29; Conservative 24; Mismatches 43; Indels 26; Gaps 4;

QY 1 LOEQORDLEQRKADTKKLE-----RKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSL 55
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 ISKQEAWEKGDLETPARMLKADIDAKKEYEEL--EQLSKETESAS-----P 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 56 HOSLPODNRGNSRDSKEISIIKTNRRESITTNVEGRDIIHGHLSEKKDGSIKPEOKED 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 KODEPPKDEGAVEDNKNGSGESENKPSDDPEEGTSD-----EKKDPADPKPDCKPE 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 116 KS 117
   : :
Db 114 ET 115

RESULT 3
US-10-821-234-899
;; Sequence 899, Application US/10821234
;; Publication No. US20050255114A1
;; GENERAL INFORMATION:
;; APPLICANT: Labat, Ivan
;; APPLICANT: Stache-Cralin, Birgit
;; APPLICANT: Andarmani, Susan
;; APPLICANT: Tang, Y. Tom
;; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;; FILE REFERENCE: 821A
;; CURRENT APPLICATION NUMBER: US/10/821,234
;; CURRENT FILING DATE: 2004-04-07
;; PRIOR APPLICATION NUMBER: US 60/462,047
;; PRIOR FILING DATE: 2003-04-07
;; NUMBER OF SEQ ID NOS: 1704
;; SOFTWARE: PC_SEQ_genes Version 1.0
;; SEQ ID NO 899
;; LENGTH: 2432
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-821-234-899

Query Match          13.2%; Score 79.5; DB 6; Length 2432;
Best Local Similarity 26.6%; Pred. No. 6.1;
Matches 29; Conservative 17; Mismatches 38; Indels 25; Gaps 5;

QY 17 KULEKKKE-----HGDIIADLYGRLEIPALPSPENERGYIIPHOSSLPO---DNRGNS 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1746 KMERUTSIRAGIEGFLASDV-GR-----DMSASPVVSSMPEASSSSE 1792
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 69 RDSKEISIIKTNRRESITTNVEGRDIIHGHLSEKKDGSIKPEOKEDS 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1793 KODYELFVVKDTHSEKKN-----KNRDKGEKKKRDSSLRSRKRSKS 1837
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
US-10-793-626-2268
;; Sequence 2268, Application US/10793626
;; Publication No. US20050255478A1
;; GENERAL INFORMATION:
```

```
;; APPLICANT: KIMMERLY, WILLIAM JOHN
;; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;; FILE REFERENCE: PU3480US
;; CURRENT APPLICATION NUMBER: US/10/793,626
;; CURRENT FILING DATE: 2004-03-04
;; PRIOR APPLICATION NUMBER: 60/164,258
;; PRIOR FILING DATE: 1999-11-09
;; NUMBER OF SEQ ID NOS: 4472
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2268
;; LENGTH: 140
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: synthetic
;; OTHER INFORMATION: amino acid sequence
US-10-793-626-2268

Query Match          13.1%; Score 79; DB 6; Length 140;
Best Local Similarity 22.4%; Pred. No. 0.2;
Matches 28; Conservative 30; Mismatches 35; Indels 32; Gaps 6;

QY 1 LOEQORDLEQRKADTKKLEKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSL 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 LKKNKDLKKQKKLQEQKDLKKHKDPSLQEDVN-----DLPAKS-----TSR 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 PODNRGNSRDSKEISIIKT---NRRESITTNVEGRDII---KGHLSEKKDGSIKPEQ 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 DKXNKN-HDAKESKSDNSTSANHDDQTNKIKSNQDHDSSQSKPHTQO-----KPSQ 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 113 KEDKS 117
   : :
Db 130 NDRKN 134

RESULT 5
US-10-793-626-2438
;; Sequence 2438, Application US/10793626
;; Publication No. US20050255478A1
;; GENERAL INFORMATION:
;; APPLICANT: KIMMERLY, WILLIAM JOHN
;; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;; FILE REFERENCE: PU3480US
;; CURRENT APPLICATION NUMBER: US/10/793,626
;; CURRENT FILING DATE: 2004-03-04
;; PRIOR APPLICATION NUMBER: 60/164,258
;; PRIOR FILING DATE: 1999-11-09
;; NUMBER OF SEQ ID NOS: 4472
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2438
;; LENGTH: 140
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: synthetic
;; OTHER INFORMATION: amino acid sequence
US-10-793-626-2438

Query Match          13.1%; Score 79; DB 6; Length 140;
Best Local Similarity 22.4%; Pred. No. 0.2;
Matches 28; Conservative 30; Mismatches 35; Indels 32; Gaps 6;

QY 1 LOEQORDLEQRKADTKKLEKKEHGDIADLYGRLEIPALPSPENERGYIIPHOSSL 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 LKKNKDLKKQKKLQEQKDLKKHKDPSLQEDVN-----DLPAKS-----TSR 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 PODNRGNSRDSKEISIIKT---NRRESITTNVEGRDII---KGHLSEKKDGSIKPEQ 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 DKXNKN-HDAKESKSDNSTSANHDDQTNKIKSNQDHDSSQSKPHTQO-----KPSQ 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 113 KEDKS 117
   : :
Db 130 NDRKN 134
```



```
RESULT 9
US-10-793-626-2808
; Sequence 2808, Application US/107933626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2808
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2808

Query Match          12.0%; Score 72.5; DB 6; Length 491;
Best Local Similarity 21.9%; Pred. No. 3.9;
Matches 23; Conservative 24; Mismatches 37; Indels 21; Gaps 6;

Oy 7 DLEORADTKKXLEKKEHGDILAEGLYGRLEIPALPESENERYGYPHOSLPOD--N 64
Db 363 DLELRNVDDISGLYLNKQKSMVD---SMYSLSV-----RGDYGYP--NNAIYLPDGLN 412
Oy 65 RGNRSRDKSEISIEKTN-----RESITNV--EGRRDIHKGL 100
Db 413 RGNVSNYTKVTLQIKELNTTFGDKRGQVYTNELNLSKXDIPNSYI 457

RESULT 10
US-10-467-657-5486
; Sequence 5486, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5486
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5486

Query Match          11.9%; Score 71.5; DB 6; Length 212;
Best Local Similarity 23.2%; Pred. No. 1.7;
Matches 35; Conservative 21; Mismatches 38; Indels 57; Gaps 8;

Oy 14 DTKNLERKKE-----HGDLA--EDLYGRLEIPAI 42
Db 56 DSKLSEKRDALAEWKQAVAMHVAASTPEEIASLNIHLATMLAKRAVYG----- 108
Oy 43 ELPSNER-----GYIIPHOSLPOD--RGNR--DSKEISIEKTNRESITTVNVEGR- 93
Db 109 -LAARPEKFIIDGNRIPEHLGIPEAEVAVVGDSKTIIEISASVLAKTARDAEVYALAQRRP 167

Query Match          11.8%; Score 71.5; DB 6; Length 212;
Best Local Similarity 23.2%; Pred. No. 1.7;
Matches 35; Conservative 21; Mismatches 38; Indels 57; Gaps 8;
```

```
Oy 94 ----DIHKG-----HLEKKDGSIKPEOKED 115
Db 168 QYGFDKHKGVTGKHLEALQYGVLPBHRRD 198

RESULT 11
US-10-467-657-7404
; Sequence 7404, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7404
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7404

Query Match          11.9%; Score 71.5; DB 6; Length 212;
Best Local Similarity 22.2%; Pred. No. 1.7;
Matches 35; Conservative 21; Mismatches 38; Indels 57; Gaps 8;

Oy 14 DTKNLERKKE-----HGDLA--EDLYGRLEIPAI 42
Db 56 DSKLSEKRDALAEWKQAVAMHVAASTPEEIASLNIHLATMLAKRAVYG----- 108
Oy 43 ELPSNER-----GYIIPHOSLPOD--RGNR--DSKEISIEKTNRESITTVNVEGR- 93
Db 109 -LAARPEKFIIDGNRIPEHLGIPEAEVAVVGDSKTIIEISASVLAKTARDAEVYALAQRRP 167
Oy 94 ----DIHKG-----HLEKKDGSIKPEOKED 115
Db 168 QYGFDKHKGVTGKHLEALQYGVLPBHRRD 198

RESULT 12
US-10-467-657-7710
; Sequence 7710, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7710
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7710

Query Match          11.8%; Score 71; DB 6; Length 328;
Best Local Similarity 30.9%; Pred. No. 3.3;
```

Matches 25; Conservative 15; Mismatches 35; Indels 6; Gaps 3;
 QY 35 GLEIPALPSENERGY-IPHOSSLPODNR--GNSRDSKISIIETKTRNSITTVNG 91
 Db 137 GQIE---MEIKYEGGQGVSGRQYVRDENRQIGAIQLDASFSPISRVSEVEPARVQ 193
 QY 92 RRDINKHLEKKDGSIKPEQ 112
 Db 194 RTDLKLVLDIETDGSIDPEE 214

RESULT 13
 US-11-060-914-2
 ; Sequence 2, Application US/11060914
 ; Publication No. US20050261184A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaufman, Paul L
 ; APPLICANT: Kaufman, Benjamin
 ; APPLICANT: Berhnadeky, Alexander
 ; APPLICANT: Berhnadeky, Teresa
 ; TITLE OF INVENTION: Method for Treating Glaucoma
 ; FILE REFERENCE: 960296.00128
 ; CURRENT APPLICATION NUMBER: US/11/060,914
 ; PRIOR FILING DATE: 2005-02-18
 ; PRIOR APPLICATION NUMBER: US 60/545,722
 ; PRIOR FILING DATE: 2004-02-18
 ; PRIOR APPLICATION NUMBER: US 60/545,723
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2
 ; LENGTH: 793
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-060-914-2

Query Match 11.4%; Score 69; DB 7; Length 793;
 Best Local Similarity 22.4%; Pred. No. 15;
 Matches 28; Conservative 34; Mismatches 39; Indels 24; Gaps 6;
 QY 1 LOEQORDLEQKAD-----TKNLERKKEHGDILAEPLYGRLEIPALPSENE 49
 Db 579 LEEERRRKQOEADRLREBEERKRLKEIERRR-----AEAAEKQKMPBGL--SDDK 631
 QY 50 RGY--YIPHOSSLPODNRGN--SRDSKISIIETKTRNSITTVNGRDIHKHLEKKD 105
 Db 632 KPPKCTPPGSSSLKIEBRAEFLLKSVQKSGVSTHQAALVSKIDRLQGYTSAIGTK- 690
 QY 106 GSIXP 110
 Db 691 -SAKP 694

RESULT 14
 US-11-060-914-4
 ; Sequence 4, Application US/11060914
 ; Publication No. US20050261184A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaufman, Paul L
 ; APPLICANT: Kaufman, Benjamin
 ; APPLICANT: Berhnadeky, Alexander
 ; APPLICANT: Berhnadeky, Teresa
 ; TITLE OF INVENTION: Method for Treating Glaucoma
 ; FILE REFERENCE: 960296.00128
 ; CURRENT APPLICATION NUMBER: US/11/060,914
 ; PRIOR FILING DATE: 2005-02-18
 ; PRIOR APPLICATION NUMBER: US 60/545,722
 ; PRIOR FILING DATE: 2004-02-18
 ; PRIOR APPLICATION NUMBER: US 60/545,723
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4

LENGTH: 531
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-11-060-914-4
 Query Match 11.3%; Score 68; DB 7; Length 531;
 Best Local Similarity 22.0%; Pred. No. 11;
 Matches 29; Conservative 31; Mismatches 44; Indels 28; Gaps 6;
 QY 1 LOEQORDLEQKADTKKLERKKEHGDILAEPLYGRLEIPALPSENERGYIPIHQ 57
 Db 91 LERLARREBRQKRLQELERQKEDPTITD---GSLVPSRMMQNSAENS----- 139
 QY 58 SSLPODNGNSR---DSKEISIIETKTRNSITTVNGRDIHKHLEKK-----D 105
 Db 140 -TASGEKESRSGGVMEETVITTSYQKSYQDAEDKKKEKEEEREBEELKGNLGE 198
 QY 106 GSIXPEQ-KEDX 116
 Db 199 NOIKDEKIKDKX 210

RESULT 15
 US-10-467-657-1336
 ; Sequence 1336, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SPA
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MANSIGNI Vega
 ; APPLICANT: MONACI Elisabetha
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; PRIOR FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIOR FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqMan99, version 1.04
 ; SEQ ID NO 1336
 ; LENGTH: 369
 ; TYPE: PRT
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-1336

Query Match 11.1%; Score 67; DB 6; Length 369;
 Best Local Similarity 27.4%; Pred. No. 9;
 Matches 20; Conservative 14; Mismatches 29; Indels 10; Gaps 1;
 QY 40 PALPSENERGYIIPHOSSLPODNRGNSRD-----SKEISIIETKTRNSITTV 89
 Db 163 PGHTVSEYKRGGVSLIAVYQDNGKAKDIALSYAANGGTGKGVIEETTFRETTDL 222
 QY 90 EGRDINKHLEEE 102
 Db 223 FGEQAVLCGVAB 235

Search completed: December 7, 2005, 07:31:06
 Job time : 2.70208 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM proteinh - nucleic search, using frame_plus_pzn model

Run on: December 7, 2005, 07:30:46 ; Search time 2073.85 Seconds
(without alignments)
2639.584 Million cell updates/sec

Title: US-09-837-344-20
Perfect score: 603
Sequence: 1 LQEQQRDLERKADTKKLE.....GHLEKKDSGSIKEQKEDKS 117

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ pzn.model -DEV=xlh
-O/cgnt2.1/USPTO.spool.h/US09837344/runat.05122005.133532.750/app.query.fasta.1.718
-DB=EST -OPMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt -LIST=45
-DOCALLIGN=200 -THR_SCORE=pctc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pctc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09837344 @CGN 1.1 6731 @runat.05122005.133532.750 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTPUT -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_ests1.*
2: gb_ests2.*
3: gb_ests3.*
4: gb_hnc.*
5: gb_hnc4.*
6: gb_ests5.*
7: gb_ests6.*
8: gb_ests7.*
9: gb_gss81.*
10: gb_gss82.*
11: gb_gss83.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	65.7	317	10	T09652 0224m3 gmbp
2	103.5	17.2	1782	2	BF133651 601778550
3	99.5	16.5	2131	10	AG332417 Mus muscu
4	98	16.3	589	7	CV271474 WS0154.B2
5	97.5	16.2	605	6	CA082727 SCBPAM202
6	96	15.9	684	8	DR983403 JGI AOKFI
7	96	15.9	1042	3	BI559491 603252811

8	96	15.9	1082	2	BG497835	BG497835 602543125
9	95	15.8	1714	10	AG390929	AG390929 Mus muscu
10	94.5	15.7	667	6	CA083387	CA083387 SCBPAM201
11	94.5	15.7	676	6	CA160204	CA160204 SCJFR23C0
12	94.5	15.7	720	8	CX328033	CX328033 JGI_X2T67
13	94.5	15.7	974	6	CA234102	CA234102 SCBPAM225
14	94.5	15.7	1665	2	BG284329	BG284329 602408426
15	94	15.6	691	2	BG832038	BG832038 602765038
16	94	15.6	920	5	BH132586	BH132586 ENTCO242TR
17	94	15.6	1502	5	BU907774	BU907774 AGENCOURT
18	93	15.4	1218	10	CL080889	CL080889 CH216-160
19	92.5	15.3	598	10	CE688248	CE688248 tigr-gss-
20	92.5	15.3	621	10	CM816235	CM816235 pinea.259
21	92.5	15.3	1118	3	BO073445	BO073445 AGENCOURT
22	92.5	15.3	1595	10	AG365410	AG365410 Mus muscu
23	92	15.3	575	3	BM110778	BM110778 EST558314
24	92	15.3	647	5	BU295065	BU295065 603741656
25	92	15.3	959	2	B1114982	B1114982 602862806
26	91.5	15.2	386	9	AZ164328	AZ164328 SP_0075_B
27	91.5	15.2	663	3	B1068434	B1068434 B1068434
28	91.5	15.2	803	3	BP704622	BP704622 BP704622
29	91.5	15.2	1211	2	BF037585	BF037585 601461152
30	91.5	15.2	1340	10	AG334065	AG334065 Mus muscu
31	91.5	15.2	1513	10	AG453925	AG453925 Mus muscu
32	91	15.1	626	1	AL681532	AL681532 AL681532
33	91	15.1	787	8	DR849993	DR849993 JGI_CABE1
34	91	15.1	797	8	CX940605	CX940605 JGI_CAO7
35	91	15.1	811	3	B1526650	B1526650 602925809
36	91	15.1	847	8	DR873626	DR873626 JGI_CABH6
37	91	15.1	868	8	CX472102	CX472102 JGI_X2G53
38	91	15.1	910	8	CX849277	CX849277 JGI_CAL6
39	91	15.1	1505	10	AG429823	AG429823 Mus muscu
40	90.5	15.0	684	7	CO131269	CO131269 GR_EB43H
41	90.5	15.0	773	3	BU077405	BU077405 BU077405
42	90.5	15.0	816	7	CO113362	CO113362 GR_EB013
43	90.5	15.0	907	2	BG394138	BG394138 602456317
44	90.5	15.0	1151	6	CD496499	CD496499 CDA22-F06
45	90	14.9	487	3	BM249657	BM249657 K0836H12-

ALIGNMENTS

RESULT 1
LOCUS T09652 317 bp DNA linear GSS 05-DEC-2000
DEFINITION 0224m3 gmbpFHB3.1, G. Roman Reddy Plasmodium falciparum genomic
T09652 clone 0224m, genomic survey sequence.
ACCESSION T09652
VERSION T09652.1 GI:319484
KEYWORDS GSS.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
REFERENCE 1 (bases 1 to 317)
AUTHORS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
TITLE Reddy,G.R., Chakrabarti,D., Schuster,S.M., Perl,R.U., Almira,E.C.
and Dame,J.B.
JOURNAL Gene sequence tags from Plasmodium falciparum genomic DNA fragments
PUBMED prepared by the genebase activity of mung bean nuclease
8234327 Proc. Natl. Acad. Sci. U.S.A. 90, 9867-9871 (1993)
COMMENT Other GSSs: 0224m7
Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.ufl.edu
Seq primer: T3
Class: Shotgun.
Location/Qualifiers
1..317
/organism="Plasmodium falciparum"

FEATURES

source

/mol_type="genomic DNA"
/db_xref="taxon:5833"
/clone="0224m"
/lab_host="E. coli XL1-Blue"
/clone_lib="gmbpFHB3.1, G. Roman Reddy"
/note="Vector: pBluescript SK(+); Genomic DNA, from
asynchronous blood stage parasites of the cloned Honduran
HB3 isolate cultured in vitro, was digested with mung bean
nuclease in the presence of 30% formamide at 50°C
(Venrick, K.D., Imbercki, R.B., and McCutchan, T.F. 1988.
Nucleic Acids Research 16:6883-6896). The ends of the
fragments were ligated using T4 DNA polymerase, and the
fragments were ligated to EcoR V-cleaved and
dephosphorylated pBluescript SK(+). Recombinant plasmids
transformed E. coli XL1-Blue."

ORIGIN

Alignment Scores:
Pred. No.: 1,486-36 Length: 317
Score: 396.00 Matches: 77
Percent Similarity: 96.25% Conservative: 0
Best Local Similarity: 96.25% Mismatches: 3
Query Match: 65.67% Indels: 0
DB: 10 Gaps: 0

US-09-837-344-20 (1-117) x T09652 (1-317)

QY 38 GtluileproAlilEgluleuProserGluangluArglyTyrtleProhiegln 57
|||
DB 1 GAATCGAAGCTATAGAACTTCATCAGAGATGAACGTGATATATACCATCA 60
QY 58 SerSerleuProGlnAaPaaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 77
61 TCTTCTTATACCTCAGAGCAACAGAGGATAGATTCAGAGAAATATCTATAA 120
QY 78 GtluYtThAaAaAaGtluSerleThrThAaAaAaAaAaAaAaAaAaAaAaAa 97
121 GAAAAAACAATAGAACTTATTACAAATAATGTTGAAGAGCAAGGATATACATA 180
DB 98 GtluHleuGtluGtluYtAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 117
181 GGACATCTTGAAGAAAGAAAGTGTTCATTAATAACAGTACAAAGAAATATCT 240

RESULT 2
BF133651 1782 bp mRNA linear EST 24-OCT-2000
LOCUS 601778550P1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4006597 5',
DEFINITION mRNA sequence.
ACCESSION BF133651
VERSION BF133651.1 GI:10972691
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM9238 row: c column: 14
High quality sequence stop: 550.
Location/Qualifiers
1..1782

FEATURES
SOURCE

/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:4006597"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu30"
/note="Organ: lung; Vector: pCMV-SPORT6, Site 1: NciI,
Site 2: SalI; transgenic model MN1-1, expression driven by
Mtv1-Ltr enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:
Pred. No.: 0,319 Length: 1782
Score: 103.50 Matches: 31
Percent Similarity: 46.09% Conservative: 22
Best Local Similarity: 26.96% Mismatches: 39
Query Match: 17.16% Indels: 23
DB: 2 Gaps: 3

US-09-837-344-20 (1-117) x BF133651 (1-1782)

QY 2 GtngtuglInglnAaA 21
|||
DB 1194 CAAGAGCGAGACAGCGAGAGGAGAGACAAAGAAAGAAAGAAAGAGAGAGAG 1253
QY 22 TysTysGtluHleuAa 41
1254 AGACAAGAAACAGACAGAGAAACAAAGAGAGAGCAACAAAGAGATAGACCAAGAG 1313
QY 42 lEgluleuProserGluangluArglyTyrtleProhieglnSerSerleuPro 61
1314 GCGCAACAGCAACAGAAACAGAAAGCA----- 1343
DB 62 GlnAaPaaAa 81
1344 ---GAGAAAGAAAGAAAGAAAGAG-----AAGAGAGAGAGC 1379
QY 82 ArgGtluSerleThrThAa 101
1380 AGAGAAAGAAACAGAGCGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1424
QY 102 GtluYtAa 116
DB 1425 GAGAGCAAGAAAGAAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1469

RESULT 3
AG332417/c 2131 bp DNA linear GSS 18-DEC-2004
LOCUS Mus musculus molossinus DNA, clone: M5hg01-122J15.T7, genomic survey
DEFINITION sequence.
ACCESSION AG332417
VERSION AG332417.1 GI:47905727
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Abe, K., Noguchi, H., Tagawa, K., Yuzurika, M., Toyoda, A., Kojima, T.,
Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K., and
Shioishi, T.
TITLE Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
JOURNAL Genome Res. 14 (12), 2439-2447 (2004)
PUBMED 15574823
REFERENCE
2 (bases 1 to 2131)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.

QY 22 LyslysgluhihsglyaspriileuallagluapleuTYRGLYArgleuGluileprola 41
 Db 461 GATGCTGAAAATGTTGAT-----GCAAGTATGTTTATGCCCCACAGATGACAAAAG 408
 QY 42 lileguleuProsergluaangluarglyTYRlileProhihlglnSerleu--- 60
 Db 407 AGGGA-----AAGCACAAAGATAGAAACACCGAAACGGCATCAAAAGTGCCTATGAT 354
 QY 61 -----ProgluaapnaarglyuanseraArgapserlygluileserile 76
 Db 353 GATGTAATTCTGACAGAGATGAGAAAGAGTCTTAAAAATCCGACAGCATGACGAT 294
 QY 77 lilegulythrAsnArglyuserilerThraenValgluGLYArgArgAspillehis 96
 Db 293 GACCTTAAAAATCTGAAAAGCACATACACCTGATCAGATGCTGAAAACGACGAT 234
 QY 97 LysglYhihSleuGluGluLysLysAspGlyserilerleProgluGlnLysgluAsp 115
 Db 233 AAAAGCACAAAGAGATCACCGTGATGTTCCGCAAAATGCGAGTAAATGAGAG 177
 RESULT 5
 CA082727 605 bp mRNA linear EST 23-SEP-2003
 LOCUS SCBFAM2025D06.g AM2 Saccharum officinarum cDNA clone SCBFAM2025D06
 DEFINITION 5', mRNA sequence.
 CA082727
 VERSION CA082727.1 GI:34936038
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.
 1 (bases 1 to 605)
 REFERENCE
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 The libraries that made SUCSEST
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parud@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 025 row: D column: 06
 Seq primer: 77 Promoter Primer.
 FEATURES
 source
 1..605
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCBFAM2025D06"
 /lab_host="DH10B"
 /clone_lib="AM2"
 /note="Organ: Apical meristem and tissues surrounding of
 immature plants; Vector: pSPORT1, Site 1: SalI, Site 2:
 NotI; An unidirectional cDNA library generated from
 [Apical meristem and tissues surrounding of immature
 plants]. cDNA was prepared from polyA+ mRNA using
 Superscript Plasmid System Kit (Invitrogen). The
 double-strand cDNAs were fractionated in a sepharose
 CL-2B 40cm-columns and fragments sizing between 0.8 and
 1.5 Kb were directionally cloned into the vector. Details
 of each source of RNA and library construction can be
 obtained at http://sucsest.lad.ic.unicamp.br/public"

Score: 97.50 Matches: 32
 Percent Similarity: 49.59% Conservative: 28
 Best Local Similarity: 26.45% Mismatches: 46
 Query Match: 16.17% Indels: 15
 DB: 6 Gaps: 5
 US-09-837-344-20 (1-117) x CA082727 (1-605)
 QY 2 GlnGluGlnGlnArgAspLeuGluGlnArgGlyAlaAspThrLysLysAsnLeuGluArg 21
 Db 76 GAAAAACAGAGAGAGGGGCAAAAGACAAAGAAAAAGCAAGACCGAAGATGGAACAG 135
 QY 22 LyslysgluhihsglyaspriileuallagluapleuTYRGLYArgleuGluileprola 36
 Db 136 AAGAGGGCCATCTGGAATTAAGATGATGTAACCCGACACAGAGATGTTCCATGTAAC 195
 QY 37 leuGluileprolaileguleuProsergluaangluarglyTYRlileProhis 56
 Db 196 AAGAGTTGGCAGAGAGTGCAGACTCGGGACCCGTGA-----ATCAAGCTA 243
 QY 57 GlnserleuProgluaapnaarglyuanseraArgapserlygluileserile 76
 Db 244 ACTGACTGAACACAGAGAAAGAAATGATGTAAGTTGCAAGGC----- 294
 QY 77 lilegulythrAsnArglyuserilerThraenValgluGLYArgArgAspillehis 95
 Db 295 ---AAGAAAGAAACAAAGAGAGGACCAAGGATGATCAAGAAAAAGAGACGTT 351
 QY 96 HishlysglyhihSleuGluGluLysLysAspGlyserilerleProgluGlnLysgluAsp 115
 Db 352 GGTCAAGGCGATGAAGGTGAGAAAGAAACAAAGAA--AAGAAAGAGAGAAAGAGAC 408
 QY 116 Lys 116
 Db 409 AAG 411
 RESULT 6
 DR983403 684 bp mRNA linear EST 03-AUG-2005
 LOCUS JGI_AOKF1090.rev AOKF Acropora palmata spawned eggs Acropora
 DEFINITION JGI_AOKF1090.rev AOKF Acropora palmata spawned eggs Acropora
 palmata cDNA clone AOKF1090 3', mRNA sequence.
 CA082727
 VERSION DR983403.1 GI:71778109
 KEYWORDS EST.
 SOURCE Acropora palmata
 ORGANISM Acropora palmata
 Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
 Asterozoa; Acroporidae; Acropora.
 1 (bases 1 to 684)
 REFERENCE
 Schwarz,J.A., Brokslein,P., Manohar,C., Coffroth,M.A., Samant,A.
 and Medina,M.
 Coral-Symbiodinium EST Project
 JOURNAL Unpublished (2005)
 COMMENT Other ESTs: JGI_AOKF1090.fwd
 Contact: Schwarz, JA, Medina, M.
 Evolutionary Genomics
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925-296-5823
 Email: jaszczwarz@lbl.gov
 cDNA Library Preparation: DOE Joint Genome Institute:
 http://www.jgi.doe.gov
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
 Naming Conventions: EST name is generated by the concatenation of
 the JGI Clone Id and the direction of sequencing. The suffix '.rev'
 indicates a reverse sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Poly-T: A run of 14 or more T residues at the beginning of this
 sequence has been removed.
 Plate: AOKF 0009 row: d column: 10
 High quality sequence scop: 546.
 FEATURES
 source
 1..684
 Location/Qualifiers

Alignment Scores:
 Pred. No.: 0.421 Length: 605

/organism="Acropora palmata"
/mol_type="mRNA"
/db_xref="taxon:6131"
/clone="AOKF1090"
/dev_stage="Freshly spawned eggs"
/lab_host="Electromax DH108"
/note="Vector: pDNR-LIB, Site 1: SfiI, Site 2: SfiI, The library was prepared from total RNA using the Creator SMART cDNA Library Construction Kit with the LD-PCR method to amplify the cDNA. Amplified cDNA was digested with SfiI, size selected for >400bp, and ligated into the pDNR-LIB vector."

ORIGIN

Alignment Scores:

Pred. No.: 0.737 Length: 684
Score: 96.00 Matches: 40
Percent Similarity: 43.14% Conservative: 26
Best Local Similarity: 26.14% Mismatches: 33
Query Match: 15.92% Indels: 54
DB: 8 Gaps: 8

US-09-837-344-20 (1-117) x DR983403 (1-684)

QY 1 LeuGlnGlnGlnGlnAArgPLeuGlnAArgLyAAlaAsp----- 14
DB 486 TTGCAGAGCAGCAGTGAAGAACTCGAAAGAGCGTGTGAAGTCCACTCT 427
QY 15 -----ThrylYsYsAenLeuGlu-----ArgLySlyYsGluHHisGlyAsp----- 27
DB 426 TACTTCAGAGAGAGCTTCGTAATTCGACAGACAGACAGACAGATGGCGAGAGAGA 367
QY 28 ---IleLeuAlaGluAspLeuTyrgLyArgLeuGluIleProAlaIleGluLeuProSer 46
DB 366 AAGCTCACAGACAGCAGATTGTGACAGCAAACTGAATC-----CTACATCA 319
QY 47 GLuAenGluArgGlyTyrrTyrlleProHisIleSerIleuProGlnAspAsnArgGly 66
DB 318 GAAACCGAA-----ACATTAATCCAGTTTG----- 292
QY 67 AAnSerArgAspSerLySgluIleSerIleIleGlu----- 78
DB 291 GGGAGTGAAGTGAACCTGATTTGCGTTCGGAAGAAAAAGAAAGTTCCTGTGACA 232
QY 79 ---LySerThraAsnArgIuSerIleThraAsnValGluGlyArgAsp----- 94
DB 231 ATTAAAGTCTGAGGATTAACCTGCTGATCAAGCAAACTGTAAGAGACTACAGACAG 172
QY 95 -----IleHisLySglYHisLeuGlu 101
DB 171 AAGTCGCAATGATGAAGAGAAATTAGAGACTTTTCAGAGTCGAGAAAGCGCCTTGAA 112
QY 102 GLuLyLyAspArgLySerIleLyProGluGlnLySglu 114
DB 111 AAGGAAAAAGATGGTTTCGTGCTCGGAGAAAGTCAGAGA 73
RESULT 7
BI559491 1042 bp mRNA linear EST 05-SEP-2001
LOCUS 6032581F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:525338 5',
DEFINITION mRNA sequence.
ACCESSION BI559491
VERSION BI559491.1 GI:15446805
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE 1 (bases 1 to 1042)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D., Shitaki
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1A11747 row e column: 03
High quality sequence stop: 12.
Location/Qualifiers

FEATURES

source

1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:525338"
/lab_host="DH108"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gagag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to R0F 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGRI/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 1.25 Length: 1042
Score: 96.00 Matches: 36
Percent Similarity: 41.48% Conservative: 20
Best Local Similarity: 26.67% Mismatches: 57
Query Match: 15.92% Indels: 22
DB: 3 Gaps: 5

US-09-837-344-20 (1-117) x BI559491 (1-1042)

QY 3 GlnGlnGlnAArgPLeuGlnAArgLyAAlaAspThrylYsAenLeuGluArgLyS 22
DB 474 CAGGAAAAACACCAAGCAAAAGACGAGAAACGACAAAGCAAGACAGACAA 533
QY 23 LySgluHHisGlyAsp---IleLeuAlaGluAspLeuTyrgLyArgLeuGluIleProAla 41
DB 534 AGGAGCAGAAAGCAGAGATTAAGAGAGAGAGAGAGAGAGATGAAGACGGG 593
QY 42 IleGluLeuProSerGluAenGluArgGlyTyrrTyrlleProHisIleSerIleuPro 61
DB 594 AGGGA---CCAGACAAAGACGAGCGGGAAGCAAAACAGAAAGCAAGAGA 650
QY 62 GlnAspAsnArg-----GlyAenSerArgAspSerLySgluIleSerIleIleGluLyS 79
DB 651 CAGGACAAAGAAACACACAGAAAGAAAGAGAGAAACACAAACAAAGAGCGAAAG 710
QY 80 ThraAsnArgIuSerIleThraAsnValGluGlyArgGlyAspIleHisLyS----- 97
DB 711 AGACACAGAACCCGAGCAGCAGAGAAAGAAAGAGAAAGCAAGACACCA 770
QY 98 -----GlyHis-----LeuGlu 101
DB 771 GGGAAACGAGACAGACAGAGAAAGACGACACGAGAGAGAGAGAGAGAGAGAA 830
QY 102 GLuLyLyAspArgLySerIleLyProGluGlnLySgluAspLyS 116
DB 831 GAAAGAAAGACAAAGAGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 875
RESULT 8
BG497835

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:465487"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 60"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgcgcctcggcc); Site 2: SfiI
(gggcctatggcc); 5' and 3' adaptors were used in cloning
cell line RNA.
follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and
3' adaptor sequence:
5'-ATTATGAGGCCGAGCGGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.5
kb (range 0.9-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

```

Db	714	AAACAAGAGAGAGAGAGAGTATAAGA-----GAGAAAAAAGA	752
Qy	82	ArgGluser-----lIeThrAsnValGluGlyAArgAspIle	95
Db	753	AGAGAGAGCCAGAGAGAGAAAAAGCCGAGACAGAGAAAAAGAGAAAGACAGC	812
Qy	96	HiLysGlyIstIsleuGluGlyIstLysAspGlySerIleLysProGluGlnLysGlu	114
Db	813	AAAGAGGAGAAAGAGAAAGAAAGAGAAAGGAGAGAGAGAGAGAAAGAGAGAG	869
RESULT 9			
AG390929		1714 bp	DNA
AG390929			linear
AG390929			GSS 21-DEC-2004
LOCUS			
DEFINITION			Mus musculus molossinus DNA, clone:MSMG01-207N12.TU, genomic survey
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			Mus musculus molossinus (Japanese wild mouse)
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
PRIMERS			
SEQUENCING			
LIBRARY			
VECTOR			
R SITE 1			
R SITE 2			
FEATURES			
SOURCE			
ORIGIN			
ALIGNMENT SCORES			
PRED. NO.:			
SCORE:			
PERCENT SIMILARITY:			
BEST LOCAL SIMILARITY:			
QUERY MATCH:			
DB:			


```

source
1..676
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCUFR23C08B07"
/lab_host="DH10B"
/clone_lib="R23"
/note="Organ: Shoot-root transition zone from adult
plants; Vector: pSport1; Site_1: SalI; Site_2: NotI. An
unidirectional cDNA library generated from shoot-root
transition zone from adult plants). cDNA was prepared
from polyA+ mRNA using Superscript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a Sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://succeat.lad.ic.uniscamp.br/public"

ORIGIN

Alignment Scores:
Pred. No.: 1.09 Length: 676
Score: 94.50 Matches: 31
Percent Similarity: 49.59% Conservative: 29
Best Local Similarity: 25.62% Mismatches: 46
Query Match: 15.67% Indels: 15
DB: Gaps: 5

US-09-837-344-20 (1-117) x CA160204 (1-676)

OY 2 GINGLUGLNGlnArGApLeuGluGlnArGlyAlaApThrLySAsnLeuGluArG 21
   ::::::::::: ::::::::::: :::::::::::
DB 86 GAAAAACAAAGAAAGCAAGCAAGAAAAAGAAAGACAGCGCAAGATGCAAGC 145
   ::::::::::: ::::::::::: :::::::::::
OY 22 LyAlYsGlnHis-----GlyApRlLeuAlaGluAspLeuTyRgLyArG 36
   ::::::::::: ::::::::::: :::::::::::
DB 146 AAGGAGCGCCATCTGGAGATAGAAAGATGGTGAAGCCGACAGAGATGTCATGAGAC 205
   ::::::::::: ::::::::::: :::::::::::
OY 37 LeuGluIlleProAlaIleGluLeuProSerGluAsnGluArGlyTyRtlleProHis 56
   ::::::::::: ::::::::::: :::::::::::
DB 206 AAGAGCTGGCAGAGCTGCAGACTCGGGAGCCCGTGA-----ATCAAGCTA 253
   ::::::::::: ::::::::::: :::::::::::
OY 57 GlnSerSerLeuProGlnAspAsnArGlyAsnSerArGApSerLySglulleSerile 76
   ::::::::::: ::::::::::: :::::::::::
DB 254 ACTGACACTGCAACCAACGAGAAAGAAATTGATGATGAGATTCAGAAAGC----- 304
   ::::::::::: ::::::::::: :::::::::::
OY 77 lIeGluYThrAsnArGluSerlIeThrAsnVal--GluGlyArGArGApIle 95
   ::::::::::: ::::::::::: :::::::::::
DB 305 ---AAGAAAGAAAGAAAGAGAGCAAGAGATGAATCAAGAAAAAGAAAGACGTT 361
   ::::::::::: ::::::::::: :::::::::::
OY 96 HisLyGlyHisLeuGluGluYLyAspGlySerlIeYsProGluGlnYsGluArP 115
   ::::::::::: ::::::::::: :::::::::::
DB 362 GGTCAAGCGCGATGAGAGTGAAGAAAGAACAAAGAA--AAGAAAGAGAAAGAAAGAC 418
   ::::::::::: ::::::::::: :::::::::::
OY 116 Lys 116
   |||
DB 419 AAG 421

RESULT 12
CX328033 720 bp mRNA linear EST 04-JAN-2005
LOCUS CX328033
DEFINITION JGI XZT67285.fwd NIH XGC tropTad5 Xenopus tropicalis cDNA clone
IMAGE:7786463 5', mRNA sequence.
ACCESSION CX328033
VERSION CX328033.1 GI:57064505
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 720)
Richardson,P., Lucas,S., Rokhsar,D., Dettler,J.C., Ng,D.C.,

```

```

TITLE
JOURNAL
CONTACT Richardson,P.
COMMENT
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Richard M. Harland Laboratory, University of
California, Berkeley: http://tropicalis.berkeley.edu/home
cDNA library preparation: Richard M. Harland Laboratory, University
of California, Berkeley
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: XZT 0701 row: 1 column: 21
High quality sequence stop: 668.
Location/Qualifiers
1..720
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7786463"
/tissue_type="whole embryo"
/dev_stage="Tadpole (St. 36-41)"
/lab_host="E. coli XL1-Blue derivative, Stratagene
Electroten-Blue"
/clone_lib="NIH XGC tropTad5"
/note="Vector: pCS108; Site 1: SalI; Site 2: NotI; Tadpole
library constructed by Russell B. Fletcher in R. Harland's
lab using poly A RNA and oligo dt primers (Invitrogen
Superscript Plasmid System for cDNA Synthesis and
Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted
into vector pCS108
(http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)
"

ORIGIN

Alignment Scores:
Pred. No.: 1.18 Length: 720
Score: 94.50 Matches: 30
Percent Similarity: 46.03% Conservative: 28
Best Local Similarity: 23.81% Mismatches: 35
Query Match: 15.67% Indels: 35
DB: Gaps: 4

US-09-837-344-20 (1-117) x CX328033 (1-720)

OY 1 LeuGlnGluGlnGln-----ArGApLeuGluGlnArGlyAla 13
   ::::::::::: :::::::::::
DB 253 CTAAAGAGAGAAAGAGAAAGAGTGAAGAAATTCAAGATATGCAACAAAAAGCCG 312
   ::::::::::: :::::::::::
OY 14 AspThrLySAsnLeuGluArGlyLySgluHisGlyApRlLeuAlaGluAspLeu 33
   ::::::::::: :::::::::::
DB 313 GAATCAAAACAGCGCTTGAAGAAATGGAAGCAACTTAAGAGATT----- 357
   ::::::::::: :::::::::::
OY 34 TyRgLyArGLeuGluIlleProAlaIleGluLeuProSerGluAsnGluArGlyTyR 53
   ::::::::::: :::::::::::
DB 358 -----CGTTGA----- 366
   :::::::::::
OY 54 lIleProHisGlnSerSerLeuProGln-----AspAsnArGlyAsnSerArGAp 70
   ::::::::::: :::::::::::
DB 367 -----GAAAAATCACTGACGAGAAAGCAGACTGCAAGAAAGAAATTAAGAA 417
   :::::::::::
OY 71 SerLySgluIlleSerlIeIleGluYThrAsnArGluSerlIeThrAsnValGlu 90
   :::::::::::
DB 418 CAGAAAGAAATTTGGAAAAAGAAAGAAATATATATAGCCGCGTCAATTCATGAGAACGT 477

```



```
Query Match: 15.67% Indels: 33
DB: 2 Gaps: 5
```

US-09-837-344-20 (1-117) x BG284329 (1-1665)

Oy		2	GlnGluGlnGlnGlnArgAspLeuGlutGlnLysAlaSerThrLybVleAsnLeuGluArg	21
Db		903	AAAGAACCGCAGGAACAGCACCAACAAAAGAAAAAAGAAGCCGGAAACAGAACAGACGAG	962
Oy		22	LybVleGluIuhIeGIlyAspIleLeuAlaGluAspLeuTrgIyArg-----	36
Db		963	AACAAGAGGCAGAGGGAA-----GAGAAAGAGAGCAACGAAAGCAACACAAAAGAAAG	1016
Oy		37	-----LeuGluIleProAlaIleGluLeuProSer	46
Db		1017	AATGAGAAAAACAACAGAAAAGAGAGAAAAAGMAACGAGAACGAGGCCGCAAGAA	1076
Oy		47	GluAsnGluArgGlyTrpTyrIleProHieGlnSerSerLeuProGluAspAsnArgIly	66
Db		1077	AAGAAGCAAGAAC-----AAAGAACAGAGAACAC	1103
Oy		67	AenSerArgAspSerIleGluIleSerIleIleGluLybThrAsnArgGluSerIleThr	86
Db		1104	GGAACACGACAGACGAGAAAACAG-----CGAGAAACGAAACAGAGAAAC--ACA	1148
Oy		87	ThraAsnValGluGlyArgArgAspIleHieIlybVleGlyHieSleuGluGluIlyLeaAspIly	106
Db		1149	AAGAGAACACACGAGAACAGAGAGAAAAAGAGAAAAAGAAAAAAGAAAAAAGAG	1208
Oy		107	SerIleLybProGluGlnIleGluAspIlys	116
Db		1209	CAAAAAAGAAACAGACGACAGAGACAG	1238

RESULT 15
BG832038
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

691 bp mRNA linear EST 22-MAY-2001
602765038p1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4907265 5',
mRNA sequence.
BG832038
BG832038.1 GI:14179625
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 691)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
<http://image.llnl.gov>
Plate: LLCM1811 row: k column: 10
High quality sequence set: 45.

following 5 adaptor: GGCAACGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library. |"

ORIGIN

Alignment Scores:

Pred. No.:	1, 28	Length:	691
Score:	94.00	Matches:	38
Percent Similarity:	45.19%	Conservative:	23
Best Local Similarity:	28.15%	Mismatches:	41
Query Match:	15.59%	Indels:	33
DB:	2	Gaps:	6

US-09-837-344-20 (1-117) x BG832038 (1-691)

OY		1	GlnAsGlyValLeaSerThrLysLysValn-----LeuGluArgLysLysGlu-----	24
Dd		239	AAGCGAAAGGAAGATACAAAAAACAACATACAGTGTATGAAAAGAAAGAAAAAACAA	298
OY		25	-----HisGlyAspIleLeuAlaGluAsp-----LeuTyrgLyArg 36	
Dd		299	GCATATCATAGCAAAATCATAGACAGAGATCAAACAGACAAACAACAAACACAA	358
OY		37	LeuGluIleProAlaIleGluLeuProSerGluMetGluArgGlyTyrTyrIleProHie	56
Dd		359	GTRGACATTCMAACACATAGAACACCCGAAGAAAAATAAAGA-----ATA	403
OY		57	GlnSerSerLeuProGlnAspAsnArgGluValnSer-----	68
Dd		404	AACACAGACTGTGAAGAGGACAAAAAAACAAACACACAGAAAAACATCAAAACACAC	463
OY		69	-----ArgAspSerLysGluIleSerIleIleGluIyPThrAsnArgGlu-----	83
Dd		464	MGAACGACACACACAAAGAAAGAAAGAACAAACAAACAAAGATTCAAAGACAGAAC	523
OY		84	---SerIleThrThrasnValGluGly-ArgArgAspIleHisLysGluHisLeuGluG1	102
Dd		524	TAGACGAAACACAGAAACAAAGACGATAGAAAAGAGATTCAAAAGAAAGAACAAAGACA	583
OY		102	LysLysAspArgLysSerIleLysProGluGlnLysGluAspLys 116	
Dd		584	AAATTAAAGAAACAAACATTAAGACGACAAAAAGAAAGAACAAAGA 626	

Search completed: December 7, 2005, 12:38:56
Job time : 2081.85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:28:45 ; Search time 2602.64 Seconds
(without alignment)
2555.359 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LOEQQRDLQRKADTKKNLE.....GHLEKKGGKSTPEQKEDKS 117

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xjh
-O/cgpn2.1/uspto.spool.h/US09837344/runat.05122005.133532.734/app.query.fasta.1.718
-DB=GenEmbl -QPM=fastap -SUFFIX=gc -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prc -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09837344 @CGN 1.1 5142 @runat.05122005.133532.734 -NCPU=6 -ICPU=3
-NO_MMP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_pro.*
10: gb_sta.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_hng.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603	100.0	950	6	A28740 CDNA for LS
2	603	100.0	950	6	ARI62961 Sequence
3	603	100.0	950	6	ARI78250 Sequence

4	603	100.0	1482	6	ARI62970 Sequence
5	603	100.0	1482	6	ARI78259 Sequence
6	603	100.0	1493	6	A28743 CDNA for LS
7	603	100.0	1493	6	ARI62967 Sequence
8	603	100.0	1493	6	ARI78256 Sequence
9	603	100.0	1496	6	PFLSA13
10	602	99.8	1368	11	AY751501
11	602	99.8	5970	2	PFLSA1G
12	600	99.5	1482	6	ARI62974
13	600	99.5	1482	6	ARI78263
14	597	99.0	253001	2	AE014834
15	554	91.9	843	2	PFLSA1C
16	554	91.9	843	2	PFLSA1B
17	554	91.7	843	2	PFLSA1A
18	553	91.7	843	2	PFLSA1D
19	553	91.7	843	2	PFLSA1E
20	553	91.7	843	2	PFLSA1F
21	553	91.7	843	2	PFLSA1G
22	550	91.2	843	2	PFLSA1H
23	550	91.2	843	2	PFLSA1I
24	549	91.0	843	2	PFLSA1J
25	549	91.0	843	2	PFLSA1K
26	549	91.0	843	2	PFLSA1L
27	549	91.0	843	2	PFLSA1M
28	549	91.0	843	2	PFLSA1N
29	549	91.0	843	2	PFLSA1O
30	549	91.0	843	2	PFLSA1P
31	549	91.0	843	2	PFLSA1Q
32	549	91.0	843	2	PFLSA1R
33	469	77.8	795	2	AF246996
34	224	37.1	126	2	AY090552
35	104.5	17.3	161790	14	AC022371
36	104.5	17.3	161863	8	AL606970
37	103.5	17.2	162642	5	BX908804
38	102	16.9	13097	1	AE000673
39	102	16.8	218208	14	AC073700
40	101.5	16.8	110000	15	CR382133_11
41	98.5	16.3	107112	8	AL596210
42	98.5	16.3	184515	14	AC073482
43	98.5	16.3	347050	2	PFA929351
44	97	16.1	2130	6	AR549811
45	97	16.1	112968	8	AC093833

ALIGNMENTS

RESULT 1	A28740	950 bp	DNA	linear	PAT 04-JUN-1995
LOCUS	A28740				
DEFINITION	CDNA for LSA-R-NR protein.				
ACCESSION	A28740				
VERSION	A28740.1	GI:1247512			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 950)				
JOURNAL	LIVER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF P. FALCIPARUM BEARING EPTOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES				
FEATURES	Patent: WO 921384-A 29 20-AUG-1992; Location/Qualifiers				
source	1..950				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN
Alignment Scores: 1.31e-50 length: 950
Pred. No.: 603.00 Matches: 117
Score:

VERSION AR162970.1 GI:16233435
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1482)
AUTHORS Guerin-Marchand, C. and Druilhe, P.
TITLE Peptide sequences specific for the hepatic stages of *P. falciparum*
JOURNAL Patent: US 6270771-A 41 07-AUG-2001;
FEATURES Location/Qualifiers
source 1..1482
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.:	2,096-50	Length:	1482
Score:	603.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-837-344-20 (1-117) x AR162970 (1-1482)

QY 1 LeuGlnGlnGlnGlnArgAspLeuGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
|||||
610 TTCCAGAGACGACGAAAGAGATTAGAACAAAGAGGCTGATACGAAAAAAATTTAGAA 669
DB 21 ArgLysLysGluHisGlyAspLleLeuAlaGluAspLeuTyrgLysArgLeuGluLlePro 40
|||||
670 AGAAAAAGGACGACGAGATATATTAGCAGAGGATTATATGTCGTTAGAAATACCA 729
QY 41 AlaIleGluLeuProSerGluAsnGluArgLysTyrgLysLeuProHisGlnSerSerLeu 60
|||||
730 GCTATAGAACTTCATCAGAAATGAAAGCGATATTATACCACTCAATCTTCTTTA 789
DB 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluLleSerLleIleGluLysThr 80
|||||
790 CCTCAGACAAACAGAGGAAATAGTAGAGATTCCAGAGAAATATCTATAAGAAAAACA 849
QY 81 AsnArgGluSerLleThrThrAsnValGluGlyArgArgAspLleHisLysGlyHisLeu 100
|||||
850 AATAGAGAAATCTATTACAAACAATGTTGAAGSACGAAAGGATATACATTAAGGACATCTT 909
QY 101 GluGluLysLysAspGlySerLleLysProGluGlnLysGluAspLysSer 117
|||||
910 GAAGAAAAAGAAAGATGGTTCATTAACACGAAACAAAAAGAAATTAATCT 960

RESULT 5
AR178259 1482 bp DNA linear PAT 20-APR-2002
LOCUS AR178259
DEFINITION Sequence 41 from patent US 6319502.
ACCESSION AR178259
VERSION AR178259.1 GI:20219397
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1482)
AUTHORS Guerin-Marchand, C. and Druilhe, P.
TITLE Peptide sequences specific for the hepatic stages of *P. falciparum*
JOURNAL Patent: US 6319502-A 41 20-NOV-2001;
FEATURES Location/Qualifiers
source 1..1482
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.:	2,096-50	Length:	1482
Score:	603.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-837-344-20 (1-117) x AR178259 (1-1482)

QY 1 LeuGlnGlnGlnGlnArgAspLeuGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
|||||
610 TTCCAGAGACGACGAAAGAGATTAGAACAAAGAGGCTGATACGAAAAAAATTTAGAA 669
DB 21 ArgLysLysGluHisGlyAspLleLeuAlaGluAspLeuTyrgLysArgLeuGluLlePro 40
|||||
670 AGAAAAAGGACGACGAGATATATTAGCAGAGGATTATATGTCGTTAAGAAATACCA 729
QY 41 AlaIleGluLeuProSerGluAsnGluArgLysTyrgLysLeuProHisGlnSerSerLeu 60
|||||
730 GCTATAGAACTTCATCAGAAATGAAAGCGATATTATACCACTCAATCTTCTTTA 789
DB 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluLleSerLleIleGluLysThr 80
|||||
790 CCTCAGACAAACAGAGGAAATAGTAGAGATTCCAGAGAAATATCTATAATAGAAAAACA 849
QY 81 AsnArgGluSerLleThrThrAsnValGluGlyArgArgAspLleHisLysGlyHisLeu 100
|||||
850 AATAGAGAAATCTATTACAAACAATGTTGAAGSACGAAAGGATATACATTAAGGACATCTT 909
QY 101 GluGluLysLysAspGlySerLleLysProGluGlnLysGluAspLysSer 117
|||||
910 GAAGAAAAAGAAAGATGGTTCATTAACAAACGAAACAAAAAGAAATTAATCT 960

RESULT 6
A28743 1493 bp DNA linear PAT 04-JUN-1995
LOCUS A28743
DEFINITION cDNA for LSA gene 3' (partial).
ACCESSION A28743
VERSION A28743.1 GI:1247516
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1493)
AUTHORS LIVER-STAGE-SPECIFIC PEPTIDE SEQUENCES OF *P. FALCIPARUM* BEARING
TITLE EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES
JOURNAL Patent: WO 9213884-A 32 20-AUG-1992;
FEATURES Location/Qualifiers
source 1..1493
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	2,1e-50	Length:	1493
Score:	603.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-837-344-20 (1-117) x A28743 (1-1493)

QY 1 LeuGlnGlnGlnGlnArgAspLeuGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
|||||
607 TTCCAGAGACGACGAAAGAGATTAGAACAAAGAGGCTGATACGAAAAAAATTTAGAA 666
DB 21 ArgLysLysGluHisGlyAspLleLeuAlaGluAspLeuTyrgLysArgLeuGluLlePro 40
|||||
667 AGAAAAAGGACGACGAGATATATTAGCAGAGGATTATATGTCGTTAGAAATACCA 726

QY	41	a le gu leu Pro Ser Glu Asn Glu Val Arg Gly Tyr Tyr Ile Pro His Gln Ser Ser Leu	60
Db	727	GCTATAGAACTTCCATTCAGAAAATGAACTGGATATATTATATACCATCAATCAATCTTCTTTA	786
QY	61	Pro Gln Asp Asn Arg Gly Asn Ser Arg Asp Ser Lys Glu Ile Ser Ile le Glu Lys Thr	80
Db	787	CCTCAGACAAACAGAGGGAAATGTTAGAGTTCCAGAGAAATATCTATATATAGAAAAACA	846
QY	81	Asn Arg Glu Ser Ile Thr Thr Asn Val Glu Gly Val Glu Asp Ile His Lys Gly His Leu	100
Db	847	AATAGAGAAATCTATTATCAACAATGTTGTAAGGACGAGAGGATATCATMAAGACATCTT	906
QY	101	Glu Glu Lys Lys Asp Gly Ser Ile Lys Pro Glu Glu Lys Glu Asp Lys Ser	117
Db	907	GAAAGAAAAGAAAGATGTTCAATTAACACAGAACAAAAAGAAATTAATCT	957
RESULT 7			
LOCUS	AR162967	1493 bp	DNA
DEFINITION	Sequence 38 from patent US 6270771.	linear	PAT 17-OCT-2001
ACCESSION	AR162967		
VERSION	AR162967.1	GI:16233432	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1493)		
TITLE	Guerin-Marchand,C. and Druilleh,P. Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes		
JOURNAL	Patent: US 6270771-A 38 07-AUG-2001;		
FEATURES	location/Qualifiers		
source	1..1493		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,1e-50	Length:	1493
Score:	603.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-837-344-20 (1-117) x AR162967 (1-1493)			
QY	1	Ieu Ging Lue Pro Ser Glu Asn Glu Val Arg Gly Tyr Tyr Ile Pro His Gln Ser Ser Leu	20
Db	607	TTGCAAGAGACACAAAGAGATTTAGAACAAAGAGAGGCTGATACGAAAAAATTTAGAA	666
QY	21	Arg Lys Lys Glu His Gly Asp Ile Leu Val Glu Asp Leu Tyr Gly Val Glu Lys Leu Pro	40
Db	667	AGAAAAAAGAACATGAGATATATTATGAGAGAGATTATATGTCGTTTAGAAAAATACCA	726
QY	41	Ala Ile Glu Leu Pro Ser Glu Asn Glu Val Arg Gly Tyr Tyr Ile Pro His Gln Ser Ser Leu	60
Db	727	GCTATAGAACTTCCATTCAGAAAATGAACTGGATATATTATATACCATCAATCTTCTTTA	786
QY	61	Pro Gln Asp Asn Arg Gly Asn Ser Arg Asp Ser Lys Glu Ile Ser Ile le Glu Lys Thr	80
Db	787	CCTCAGACAAACAGAGGGAAATGTTAGAGTTCCAGAGAAATATCTATATATAGAAAAACA	846
QY	81	Asn Arg Glu Ser Ile Thr Thr Asn Val Glu Gly Val Glu Asp Ile His Lys Gly His Leu	100
Db	847	AATAGAGAAATCTATTATCAACAATGTTGTAAGGACGAGAGGATATCATMAAGACATCTT	906
QY	101	Glu Glu Lys Lys Asp Gly Ser Ile Lys Pro Glu Glu Lys Glu Asp Lys Ser	117
Db	907	GAAAGAAAAGAAAGATGTTCAATTAACACAGAACAAAAAGAAATTAATCT	957
RESULT 8			

LOCUS	AR178256		1493 bp	DNA	linear	PAT 20-APR-2002
DEFINITION	Sequence 38 from patent US 6319502.					
ACCESSION	AR178256					
VERSION	AR178256.1	GI:20219394				
KEYWORDS	.	Unknown.				
SOURCE	ORGANISM	Unknown.				
REFERENCE	AUTHORS	Unclassified.				
TITLE	Guérin-Marchand,C. and Druilhe,P. Peptide sequences specific for the hepatic stages of P. falciparum bearing epitopes capable of stimulating the T lymphocytes Patent: US 6319502-A 38 20-NOV-2001;					
JOURNAL	FEATURES	Location/Qualifiers				
source	1..1493	/organism="unknown"				
ORIGIN	/mol_type="unassigned DNA"					
Alignment Scores:						
Pred. No.:	2.1e-50	Length:	1493			
Score:	603.00	Matches:	117			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			
US-09-837-344-20 (1-117) x AR178256 (1-1493)						
QY	1	LenginglugnglnghnAAGAApPLeugLgUlaArgpysAlaApThrlvSLyASnLeuGlu	20			
DB	607	TTGCAGAAGCAGCAAGAGATTTGAACCAAGGAAAGCGCTGATACGAAAAAAAAATTTGAA	666			
QY	21	ArqlyslvsglunHISglYAsplleLeuAlagLuAspleurTYrgIYArgLeuGluilePro	40			
DB	667	AGAAAAAAGAACACTGGAGATATTGTTCAGAGAGATTTAATGTGCTTGTAAGAAATACCA	726			
QY	41	AlallegluLeuProSerGluAsnGluArgLYTyrrTYrilleProOHlgInserSerleu	60			
DB	727	GCTTAGCAACTCCATCATAGAAAAGAACGTCGATATTATATACCACTCATCTTCCTTTA	786			
QY	61	ProGlnspApnaARGglYaenseFArgAspserysgluieserllelegluysrthr	80			
DB	787	CCTCAGGCAACAAGGAGGATAGTAGAGATTCCAAAGGAAATATCTATATAGAAAAACA	846			
QY	81	AsnarrgJuseSerlleThrAsnValGlUGlYArgrArpAsplleHISLySGLYHISleu	100			
DB	847	AATAGAGAACTTATTACACAAATCTTGGAAGAGCAAGGCGATATACATMAAGACATCTT	906			
QY	101	GlugluyslvysAspglySerlleLyseProGluGlnlysgluAspLyseSer	117			
DB	907	GAAGAAAAGAAAGATGTTCAATATAAACCAAGAACAAAGATTAATCT	957			
RESULT 9						
PLSA13		1496 bp	DNA	linear	INV 05-SEP-1994	
LOCUS	P.falciparum LSA-1 gene for liver stage antigen-1 (3').					
DEFINITION	Z30320					
ACCESSION	Z30320.1	GI:510183				
KEYWORDS	liver stage antigen-1; LSA-1 gene.					
SOURCE	Plasmodium falciparum (malaria parasite P. falciparum)					
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.					
REFERENCE	1 (bases 1 to 1496)					
AUTHORS	Fidock,D.A., Gros-Masse,H., Lepers,J., Brahimi,K., Benmohamed,L., Mellouk,S., Guerin-Marchand,C., Lomono,A., Rahatimatala,L., Wats,J.F., Langley,G., Rousselhon,C., Tartar,A. and Druilhe,P. The Plasmodium falciparum liver stage antigen LSA-1 is well conserved and harbors major B- and T-cell epitopes J. Immunol. (1994) In press					
TITLE	2 (bases 1 to 1496) Fidock,D.A., Gros-Masse,H., Lepers,J., Brahimi,K., Benmohamed,L.,					
JOURNAL						
REFERENCE						
AUTHORS						

Qy 101 GIUGLULYLYAASPGLYSERILEYSPROGILUGLULYGLUAAPLYSSEr 117
DB 5236 GAAGAAAGAAAGATGCTTCATATAAACAGAACAAAAAGAAAGATTAATCT 5286

RESULT 12
ARI62974
LOCUS ARI62974 1482 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 45 from patent US 6270771.
ACCESSION ARI62974
VERSION ARI62974.1 GI:16233440
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1482)
AUTHORS Guerlin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum
JOURNAL Patent: US 6270771-A 45 07-AUG-2001;
FEATURES Location/Qualifiers
source 1..1482
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 4,14e-50 Length: 1482
Score: 600.00 Matches: 116
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.15% Mismatches: 0
Query Match: 99.50% Indels: 0
Gaps: 0

US-09-837-344-20 (1-117) x ARI62974 (1-1482)

Qy 1 LeuGInGluGlnGlnArgAspLeuGluGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
Db 610 TTGCAGAGACGACGAAAGAGATTAGAACAAAGAGAGCTGATCGAAAAAAATTTAGAA 669
Qy 21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrgLysArgLeuGluIlePro 40
Db 670 AGAAAAAGAACGACATGAGATATATTAGACAGAGATTATATGCTGCTTGAATAACCA 729
Qy 41 AlaIleGluLeuProSerGluAsnGluArgLysTyrrLysIleProHisGlnSerSerLeu 60
Db 730 GCTATAGAACTTCATCAGAAATGCAACGCGATATTATATACCATCAATCTTCTTAA 789
Qy 61 ProGlnAspAsnArgLysAsnSerArgAspSerLysGluIleSerIleGluLysThr 80
Db 790 CCTCAGAGAACGACGAGGATATGAGATTCACAGGAATGCTATTAATAGAAAAACA 849
Qy 81 AsnArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHisLysGlyHisLeu 100
Db 850 AATAGAGAACTCTATTCAACAATGTTGAAGAGCGAAGGAGATTAACATAAAGACATCTT 909
Qy 101 GUGLULYLYAASPGLYSERILEYSPROGILUGLULYGLUAAPLYSSEr 117
Db 910 GAAGAAAGAAAGATGCTTCATATAAACAGAACAAAAAGAAAGATTAATCT 960

RESULT 13
ARI78263
LOCUS ARI78263 1482 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 45 from patent US 6195502.
ACCESSION ARI78263
VERSION ARI78263.1 GI:20219401
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1482)
AUTHORS Guerlin-Marchand,C. and Druilhe,P.
TITLE Peptide sequences specific for the hepatic stages of P. falciparum

JOURNAL Patent: US 6195502-A 45 20-NOV-2001;
FEATURES Location/Qualifiers
source 1..1482
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 4,14e-50 Length: 1482
Score: 600.00 Matches: 116
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.15% Mismatches: 0
Query Match: 99.50% Indels: 0
Gaps: 0

US-09-837-344-20 (1-117) x ARI78263 (1-1482)

Qy 1 LeuGInGluGlnGlnArgAspLeuGluGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
Db 610 TTGCAGAGACGACGAAAGAGATTAGAACAAAGAGAGCTGATCGAAAAAAATTTAGAA 669
Qy 21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrgLysArgLeuGluIlePro 40
Db 670 AGAAAAAGAACGACATGAGATATATTAGACAGAGATTATATGCTGCTTGAATAACCA 729
Qy 41 AlaIleGluLeuProSerGluAsnGluArgLysTyrrLysIleProHisGlnSerSerLeu 60
Db 730 GCTATAGAACTTCATCAGAAATGCAACGCGATATTATATACCATCAATCTTCTTAA 789
Qy 61 ProGlnAspAsnArgLysAsnSerArgAspSerLysGluIleSerIleGluLysThr 80
Db 790 CCTCAGAGAACGACGAGGATATGAGATTCACAGGAATGCTATTAATAGAAAAACA 849
Qy 81 AsnArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHisLysGlyHisLeu 100
Db 850 AATAGAGAACTCTATTCAACAATGTTGAAGAGCGAAGGAGATTAACATAAAGACATCTT 909
Qy 101 GUGLULYLYAASPGLYSERILEYSPROGILUGLULYGLUAAPLYSSEr 117
Db 910 GAAGAAAGAAAGATGCTTCATATAAACAGAACAAAAAGAAAGATTAATCT 960

RESULT 14
AE014834
LOCUS AE014834 253001 bp DNA linear INV 07-OCT-2002
DEFINITION Plasmodium falciparum 3D7 chromosome 10 section 6 of 7 of the complete sequence.
ACCESSION AE014834 AE014185
VERSION AE014834.1 GI:23495173
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 253001)
AUTHORS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
Gardner,M.J., Hall,N., Pung,E., White,O., Berriman,M., Hyman,R.W., Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyes,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Angiuoli,S., Pereira,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Valdivia,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.D., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
PUBMED 12368864
REFERENCES 2 (bases 1 to 253001)
AUTHORS Gardner,M.J.
TITLE Direct Submision
JOURNAL Submitted (13-SEP-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
FEATURES Location/Qualifiers

[illegible]

	repeat_region	/note="AT rich"	/rpe_type=tandem	complement(4670, .4775)	
	repeat_region	/rpe_type=tandem	/rpe_unite="(ca)n"	4675..4768	
	repeat_region	/rpe_type=tandem	4870..4912		
	repeat_region	/rpe_type=tandem	complement(4870, .4912)		
	repeat_region	/rpe_type=tandem	/rpe_unite="(ca)n"		
	repeat_region	complement(4918, .5000)	/rpe_type=tandem		
	gene	/rpe_unite="(cctaaaa)n"	<5162..>8973		
	mrna	/locus_tag="PF10_0308"	join(<5162, .5350, .5488, .5535, .5617, .5802, .5902, .5981, .6099, .6162, .6254, .6349, .6463, .6540, .6727, .>8973)		
	CDS	/locus_tag="PF10_0308"	join(5162, .5350, .5488, .5535, .5617, .5802, .5902, .5981, .6099, .6162, .6254, .6349, .6463, .6540, .6727, .8973)		
		/locus_tag="PF10_0308"	/codon_start=1		
		/product="hypothetical protein"			

Alignment Scores:	Pred. NO.:	1.79e-47	Length:	253001
Score:	597.00	Matches:	115	
Percent Similarity:	99.15%	Conservative:	1	
Best Local Similarity:	98.29%	Mismatches:	1	
Query Match:	99.00%	Indels:	0	
DB:	2	Gaps:	0	

US-09-837-344-20 (1-117) x AE014834 (1-253001)	Oy	1	LeuGInGIuGIInGlnAArgApIeuGIuGInAGlyAlaAspThrLySylAsnLeuGIu	20
	Dd	180455	TTCGAAGAGCACCACAAAGAAATTGAGAACAAAGAGCGCTGATACGAAAAATTGAGA	180514
	Oy	21	ArGLySLySGLIHlAGLyAspRIleLauAlAGLyAspLeuTyGlyVArgLeuGIuIlePro	40
	Dd	180515	AGAAAAGGAAACAATGGAGATGTATTATGCAGAGAGTTTATATGTCGTTTAGAAATRCA	180574
	Oy	41	AlAlleGIuLeuProSeRGluAsnGIuARgLyTYrTYrIIleProHIsenSerLeu	60
	Dd	180575	GCTATGAACTCCATCACAAATGAAACGCGATATTATATATACCATCATCTCTTTA	180634
	Oy	61	ProGInAspAsnArgLyAsnSerArGaSPserLySGluIleSerIleIleGIuLythr	80
	Dd	180635	CCTCAGGACCAACGAGGGAATGTAGAGATTCCAGAGAAATATCTTAATAGAAAAATAC	180694
	Oy	81	AenArgGluSerIleThrThraenValGIuGIYArgArGaSpRIleHISylSGLYHISLeu	100
	Dd	180695	AATAGAGATCTTAATACCAAAATGTGAAAGCGAAGGATATACATAAAGACATCTT	180754
	Oy	101	GIuGIuLySLyAspGIYSerIleLySProGIuGIuLySGluAspLySer	117
	Dd	180755	CAAGAAAGAAAGATCGTTCATTAATAACCAAGAACAAAGAAAGATTAATCT	180805

RESULT 15	PPALSAC	843 bp	DNA	linear	INV 19-AUG-1995
LOCUS	PPALSAC	Plasmodium falciparum (clones KEN1-21p, KEN1-40h, KEN1-56b,			
DEFINITION	PNIG1-28a, PNIG1-28c, BRA1-19c)	liver stage-specific antigen 1			
ACCESSION	L40885	(LRA-1) gene, 3' end of cds.			
VERSION	L40885.1	GI:725278			
KEYWORDS	exoenzyme, cytochrome b, liver stage-specific antigen, liver stage-specific antigen 1, malaria.				
SOURCE	Plasmodium falciparum (malaria parasite P. falciparum)				
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT

1 (bases 1 to 843)
Yang, C., Shi, Y. P., Udhayakumar, V., Alpers, M. P., Povoa, M. M.,
Hawley, W. A., Collins, W. E. and Lal, A. A.
Sequence variations in the non-repetitive regions of the liver
stage-specific antigen-1 (LSA-1) of *Plasmodium falciparum* from
field isolates
Mol. Biochem. Parasitol. 71 (2), 291-294 (1995)
7477115

Original source text: *Plasmodium falciparum* (individual isolate
Kenyan) (clone: KEN1-56b) blood stage parasite DNA; *Plasmodium*
falciparum (individual isolate Kenyan) (clone: PNG1-28a) blood
stage parasite DNA; *Plasmodium falciparum* (individual isolate
Kenyan) (clone: PNG1-28c) blood stage parasite DNA; *Plasmodium*
falciparum (individual isolate Kenyan) (clone: KEN1-21b) blood
stage parasite DNA; *Plasmodium falciparum* (individual isolate
Kenyan) (clone: KEN1-40h) blood stage parasite DNA; and *Plasmodium*
falciparum (individual isolate Kenyan) (clone: BR1-19e) blood
stage parasite DNA.
Clone name abbreviations: KEN for Kenya; PNG for Papua New Guinea
and BRA for Brazil. For each clone, the number following the 'K',
PNG or BRA represents group number, the number following the '-',
denotes the patient's blood sample number and the lower case letter
represents recombinant clones from the corresponding isolates.
Location/Qualifiers
1. 843
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/isolate="Kenyan"
/db_xref="taxon:5833"
/clone="KEN1-56b"
/clone="PNG1-28a"
/clone="PNG1-28c"
/clone="KEN1-21b"
/clone="KEN1-40h"
/clone="BR1-19e"
/dev_stage="blood stage parasite"
1. 843
/gene="LSA-1"
<1. 843
/gene="LSA-1"
/codon_start=1
/product="liver stage-specific antigen 1"
/protein_id="AA859231.1"
/db_xref="GI:950664"
/translation="RKADITKKNLEKKKHEGDDILAEDYGRLEPALELBENERGYIT
PHOSFQPDNRGSRDSKSIIEKTRRESITTNVEGRDIIHGHELEKRGDSIKPEQ
KEDSADIIQNHLETETVNI SPDVDFQISYDELSAEDSDLI DEEDDEDLDEKPIV
QYNDFOENENIGITKELEDI EKKNEMDDDEGIEKSESELSSEKIKKGVKYEYTKDN
NPKNDKSLYDENHKYKIKYKDKQYNNKEKERIKSLFHFIDQDNEIQVDELSIDITKY
PKKL"
228
/gene="LSA-1"
/note="clone PNG1-28a"
/replace="t"

variation

ORIGIN

Alignment Scores:

Pred. No.:	8,45e-46	Length:	843
Score:	554.00	Matches:	107
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	91.87%	Indels:	0
DB:	2	Gaps:	0

US-09-837-344-20 (1-117) x PFALSAC (1-843)

QY 11 ArgLySaLaSePTTrLyElYsAsnLeUGlUArgLySLeGluHtIeGlyAaPILeUaLa 30

Db 1 AGGAAGCGTGATAGCAAAAAAATTGTGAAAGAAAAAGCAATGAGATATATTAGCA 60

QY 31 GlnuApleUeUryrGlyAYsGLeUGuILeProlaTlIeGluLeuPProSeRGluAsnGluArg 50

Db 61 GAGGATTATATATGGCGTTTAGAAATACCACTTATAGACTTCATGAGAAATGAAACGT 120

```

OY 51 G|Y|T|Y|T|Y|T|P|H|E|P|R|H|S|G|L|N|S|E|S|E|U|P|R|O|G|I|N|A|S|P|A|S|A|G|L|Y|A|S|N|S|E|R|A|G|A|P| 70
Db 121 G|G|A|T|T|T|A|T|A|T|A|C|C|A|C|A|T|C|T|T|T|A|C|C|T|C|A|G|A|C|A|A|C|A|G|A|G|G|A|T|A|G|T|A|G|A|T| 180
OY 71 S|E|R|Y|S|G|I|U|L|E|S|E|R|I|E|I|L|E|G|I|U|L|Y|S|T|H|R|A|S|N|A|R|G|I|U|S|E|R|I|E|T|H|R|T|H|R|A|S|N|V|A|I|G|I|U| 90
Db 181 T|C|C|A|A|G|G|A|A|T|A|T|C|T|A|T|A|T|A|G|A|A|A|A|A|C|A|A|T|A|G|A|A|T|C|T|A|T|T|A|C|A|C|A|A|T|G|T|G|A| 240
OY 91 G|L|Y|A|R|G|A|R|G|A|P|I|E|H|I|S|L|Y|S|G|L|Y|H|I|S|L|E|U|G|I|U|L|Y|S|L|Y|A|P|G|L|Y|S|E|R|I|E|L|Y|S|P|R|O| 110
Db 241 G|G|A|C|C|A|A|G|G|A|T|A|T|A|C|A|T|A|A|A|G|A|C|A|T|C|T|G|A|A|G|A|A|A|G|A|A|G|A|T|G|T|T|C|A|T|A|A|A|A|C|C|A| 300
OY 111 G|L|U|G|I|N|L|Y|S|G|L|U|A|P|L|Y|S|E|R| 117
Db 301 G|A|A|C|A|A|A|A|A|G|A|G|A|T|A|A|A|T|C|T| 321

```

Search completed: December 7, 2005, 10:31:35
 Job time : 2652.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:22:20 ; Search time 306.146 Seconds
(without alignments)
2547.054 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LQEQKRDLEQKADTKKLE.....GHLEKKDSIKPEKEDKS 117

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-O=/cgn2.1/USPTO.spool.h/US09837344/runat.05122005.133531.726/app.query.fasta_1.718
-DB=an.GeneSeq -OFMT=fasta -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCAL=GN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09837344 @CGN 1.1 900 @runat.05122005.133531.726 -NCPU=6 -ICPU=3
-NO MMAP -LARGEJITTER -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEF_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_GeneSeq_21.*

1: geneSeqn19808:*
2: geneSeqn19908:*
3: geneSeqn20008:*
4: geneSeqn20018:*
5: geneSeqn20028:*
6: geneSeqn20038:*
7: geneSeqn20048:*
8: geneSeqn20058:*
9: geneSeqn20068:*
10: geneSeqn20078:*
11: geneSeqn20088:*
12: geneSeqn20098:*
13: geneSeqn20108:*
14: geneSeqn20118:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603	100.0	950	2	AAQ28115
2	603	100.0	1496	2	AAQ28119
3	602	99.8	1371	12	ADQ21941
4	602	99.8	1374	12	ADQ21919

Result No.	Score	Query Match	Length	DB ID	Description
5	564	93.5	1320	2	AAQ80916
6	94	15.6	1384	14	AD240618
7	94	15.6	1385	9	ADA49438
8	94	15.6	1385	12	ADQ24116
9	90.5	15.0	969	5	AA575358
10	90.5	15.0	969	5	AA571680
11	90.5	15.0	969	5	AA592259
12	90.5	15.0	4599	2	AA727052
13	90	14.9	4080	13	ADR25559
14	89.5	14.8	696	4	AAH84231
15	89.5	14.8	699	4	AAH84230
16	89.5	14.8	717	4	AAH84229
17	89.5	14.8	774	4	AAH84228
18	89.5	14.8	819	4	AAH84227
19	89.5	14.8	1669	4	AAH84226
20	87.5	14.5	700	2	AAV88907
21	87.5	14.5	1391	3	AA257866
22	87.5	14.5	2554	6	ABA92314
23	87	14.4	903	13	AD551252
24	87	14.4	2175	12	ADN04691
25	86.5	14.3	601	13	ACN45782
26	86.5	14.3	3684	4	ABL23583
27	86	14.3	1344	4	AAH53753
28	86	14.3	1353	4	AAH53916
29	86	14.3	3001	4	AAH54449
30	86	14.3	3974	4	AAH54481
31	85.5	14.2	2298	9	ADA02680
32	85.5	14.2	2298	10	ADB72418
33	85.5	14.2	2298	10	ADE95928
34	85.5	14.2	2469	6	ABR84206
35	85.5	14.2	3440	12	AD116295
36	85.5	14.2	3477	6	ABL70013
37	85.5	14.2	3477	6	ABK64207
38	85.5	14.2	3734	12	ADL82892
39	85.5	14.2	3734	13	ADL82892
40	85.5	14.2	3739	8	AB223989
41	85.5	14.2	3740	8	ACF62748
42	85.5	14.2	3740	8	ADB20867
43	85.5	14.2	3740	9	ADA02679
44	85.5	14.2	3740	10	ADB87956
45	85.5	14.2	3740	10	ADB72417

ALIGNMENTS

RESULT 1
AAQ28115
ID AAQ28115 standard; DNA; 950 BP.
XX
AC AAQ28115;
XX
DT 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
DE P. falciparum LSA-R-NR coding sequence.
XX
KW Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
KW paludism; liver stage-specific antigen; ss.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FH repeat_region 3..629
FT /*tag= a
FT /rpt_type= TANDEM
FT repeat_unit 3..53
FT /*tag= b
XX
XX WO9213884-A1.
XX
XX 20-AUG-1992.
XX
XX 05-FEB-1992; 92WO-FR000104.

```
XX 05-FEB-1991; 91FR-00001286.
PR (INSP ) INST PASTEUR.
XX Guerimarchand C, Druilhe P;
XX WPI; 1992-299985/36.
DR P-PSDB; AAR26941.
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX Claim 17; Fig 2; 81pp; French.
XX A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
CC gIII was used to transform E.coli. The expression library was screened
CC with human antisera against antigens of all stages of P. falciparum. The
CC library was rescreened with antibodies affinity- purified on a clone
CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic LSA
CC epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
CC containing a 12-repeat region followed by a non-repeat region. Preferred
CC antigenic polypeptides of the invention are derived from the amino acid
CC sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 950 BP; 456 A; 127 C; 206 G; 161 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,996-63 Length: 950
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-837-344-20 (1-117) x AAQ28115 (1-950)
QY 1 LeuGInGluGInGlnAArgAspLeuGluGlnAArgLysAlaAspThrLysAsnLeuGlu 20
DB 600 TTGCAGAGCAGCAAGAGATTGTAGAACAGAAAGCGCTGATACGAAAAAATTTAGAA 659
QY 21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrgLysArgLeuGluIlePro 40
DB 660 AGAAAAAGGAAACATGAGATATATTAGCAGAGATTTATATGCTCGTTAGAAATACCA 719
QY 41 AlaIleGluLeuProSerGluAsnGluAArgGlyTyrgLysIleProHisGlnSerSerLeu 60
DB 720 GCTATAGAACTTCATCAGAAAAATGACGTGATATTATATACCATCAATCTTCTTTA 779
QY 61 ProGlnAAspAAspArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80
DB 780 CCTCGAGCAACAGAGGAAATGATGAGATTCCAGAAATATCTATATAGAAAAACA 839
QY 81 AsnAArgGluSerIleThrThrAsnValGluGlyAArgAspIleHisLysGlyHisLeu 100
DB 840 AATTAGAGATCTATTATCAACAAATGTTGAGAGCAAGAGATATCATTAAGGACATCTT 899
QY 101 GluGluLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 117
DB 900 GAAGAAAAGAAAGATGCTTCATTAACACAGAAACAAAAAGAAATTAATCT 950

RESULT 2
AAQ28119
ID AAQ28119 standard; DNA; 1496 BP.
XX
XX AAQ28119;
AC
XX 25-MAR-2003 (revised)
DT 08-FEB-1993 (first entry)
XX
XX P.falciparum LSA gene 3' region.
XX
```

```
KW Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
KM paludism; liver stage-specific antigen; ss.
XX Plasmodium falciparum.
OS
XX Key Location/Qualifiers
XX CDS 1..1482
FT /*cag= c
FT /product= "LSA_C-terminal_region"
FT repeat_region 37..639
FT /*cag= a
FT repeat_unit 37..187
FT /*cag= b
FT /*rpt_type= TANDEM
XX
XX W09213884-A1.
XX 20-AUG-1992.
XX
XX 05-FEB-1992; 92WO-FR000104.
XX
XX 05-FEB-1991; 91FR-00001286.
XX (INSP ) INST PASTEUR.
XX
XX Guerimarchand C, Druilhe P;
XX
XX WPI; 1992-299985/36.
XX P-PSDB; AAR26944.
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT vaccination against, treatment of and diagnosis of malaria.
XX Claim 2; Fig 8-10; 81pp; French.
XX
XX This nucleotide sequence is the 3' part of the P.falciparum liver-stage
CC specific antigen (LSA) gene. It codes for a polypeptide sequence which
CC carries a T cell epitope characteristic of a protein produced in
CC hepatocytes infected with P.falciparum. The polypeptide can be used in
CC the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 1496 BP; 716 A; 169 C; 300 G; 311 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,46-63 Length: 1496
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-837-344-20 (1-117) x AAQ28119 (1-1496)
QY 1 LeuGInGluGInGlnAArgAspLeuGluGlnAArgLysAlaAspThrLysAsnLeuGlu 20
DB 610 TTGCAGAGCAGCAAGAGATTGTAGAACAGAAAGCGCTGATACGAAAAAATTTAGAA 669
QY 21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrgLysArgLeuGluIlePro 40
DB 670 AGAAAAAGGAAACATGAGATATATTAGCAGAGATTTATATGCTCGTTAGAAATACCA 729
QY 41 AlaIleGluLeuProSerGluAsnGluAArgGlyTyrgLysIleProHisGlnSerSerLeu 60
DB 730 GCTATAGAACTTCATCAGAAAAATGAAAGTGATATTATATACCATCAATCTTCTTTA 789
QY 61 ProGlnAAspAAspArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80
DB 790 CCTCGAGCAACAGAGGAAATGATGAGATTCCAGAAATATCTATATAGAAAAACA 849
QY 81 AsnAArgGluSerIleThrThrAsnValGluGlyAArgAspIleHisLysGlyHisLeu 100
DB 850 AATTAGAGATCTATTATCAACAAATGTTGAGAGCAAGAGGATATATACATTAAGGACATCTT 909
```

QY 101 GUGLULYLYEAPGYSERILEYSPROGLUGLNLVSGIUSAPLYSSER 117
 DB 910 GAAGAAAAGAAAGATGTTCAATPAAACCAAGAACAAAGATTAATCT 960

RESULT 3

ADO21941 standard; DNA; 1371 BP.

ADO21941;

12-AUG-2004 (first entry)

USA-NRC(H) construct DNA derived from Malaria parasite USA-1.

USA-NRC; liver stage antigen; USA-1; T-cell; B-cell epitope;
 parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
 malaria parasite P. falciparum; USA-NRC(H) construct; de; gene.

Plasmodium falciparum.
 Synthetic.

Key Location/Qualifiers
 CDS 1..1371
 /tag= a
 /product= "USA-NRC(H) construct protein derived from
 Malaria parasite USA-1"

MO2004044167-A2.

27-MAY-2004.

12-NOV-2003; 2003MO-US036011.

12-NOV-2002; 2002US-0425719P.

(REED-) REED ARMY INST RES WALTER.

Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
 Barbosa A;

WPI; 2004-420309/39.

P-PSDB; ADO21942.

Recombinant USA-NRC polypeptide for use as a vaccine and a diagnostic
 agent, preferably for diagnosing malaria, comprises a liver stage antigen
 (USA-1) epitope.

Claim 16; SEQ ID NO 25; 90pp; English.

Sequence 1371 BP; 539 A; 358 C; 283 G; 191 T; 0 U; 0 Other;

The invention relates to a novel recombinant USA-NRC polypeptide
 comprising liver stage antigen (USA-1) T-cell and B-cell epitopes. USA-1
 protein is found within the parasitophorous vacuole (PV) of Plasmodium
 falciparum. The polypeptide of the invention demonstrates antimalarial
 activities and may be useful as a vaccine during immunotherapy and as a
 diagnostic agent, preferably for diagnosing malaria. The current sequence
 is that of the USA-NRC(H)Mut construct DNA of the invention. The
 construct is harmonised for expression in Escherichia coli and comprises
 the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
 parasite P. falciparum USA-1 in addition to a C-terminal His6 tag.

Alignment Scores:

Pred. No.: 6.37e-63 Length: 1371
 Score: 602.00 Matches: 116
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.15% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 12 Gaps: 0

US-09-837-344-20 (1-117) x ADO21941 (1-1371)

QY 1 LeuGInGInGInGInAArgSPLeuGInGInAArgLYSAIAspThrLYSAsnLeuGIn 20
 DB 463 CTCGAGGACCAACAGCGGCACTCGAACAAGCCCAAGCTGACAGAAAAAACTGGAA 522
 QY 21 ArgLYSLYGLuHISGLYAspIleLeuAAGIuAspLeuTYRGLYArgLeuGInIlePro 40
 DB 523 CGCAAAAAGAAACAGCGGCACTCTCGCTGAGGACCTTACGGCGCTGGAAAACTCCA 582
 QY 41 AAlleGluLeuProSerGluAsnGluAArgGLYTYRILeProHISGInSerSerLeu 60
 DB 583 GCTATCAACTCCATCCGAAAAAGAACGGGCTACTACATCCACACAGAGCAGCTTG 642
 QY 61 ProGInAspAsnAArgLYAsnSerAArgSPSerLYSGInIleSerIleGluYThr 80
 DB 643 CCACAAAGATTAATCGCGGAATCTCCCGCAAGTAAGAAATCAGATTCGAAAAAAC 702
 QY 81 AsnAArgLYSerIleThrThrAsnValGInGInYArgAArgAspIleHISLYSLYHISLeu 100
 DB 703 AACCGGAAAGCATTCACCAACGTAAGGCGCGGACATCCCAAGGCCACCTC 762
 QY 101 GUGLULYLYEAPGYSERILEYSPROGLUGLNLVSGIUSAPLYSSER 117
 DB 763 GAAGAAAAGAAAGATGTTCAATPAAACCAAGAACAAAGATTAATCT 813

RESULT 4

ADO21919 standard; DNA; 1374 BP.

ADO21919;

12-AUG-2004 (first entry)

USA-NRC(H)Mut construct DNA derived from Malaria parasite USA-1.

USA-NRC; liver stage antigen; USA-1; T-cell; B-cell epitope;
 parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
 malaria parasite P. falciparum; USA-NRC(H)Mut construct; de; gene;
 mutant.

Plasmodium falciparum.
 Synthetic.

Key Location/Qualifiers
 CDS 1..1374
 /tag= a
 /product= "USA-NRC(H)Mut construct protein derived from
 Malaria parasite USA-1"

MO2004044167-A2.

27-MAY-2004.

12-NOV-2003; 2003MO-US036011.

12-NOV-2002; 2002US-0425719P.

(REED-) REED ARMY INST RES WALTER.

Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
 Barbosa A;

WPI; 2004-420309/39.

P-PSDB; ADO21920.

Recombinant USA-NRC polypeptide for use as a vaccine and a diagnostic
 agent, preferably for diagnosing malaria, comprises a liver stage antigen
 (USA-1) epitope.

Claim 17; SEQ ID NO 3; 90pp; English.

The invention relates to a novel recombinant USA-NRC polypeptide
 comprising liver stage antigen (USA-1) T-cell and B-cell epitopes. USA-1
 protein is found within the parasitophorous vacuole (PV) of Plasmodium

CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H)Mut construct DNA of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC within the TS P. falciparum epitope and a C-terminal His6 tag.

XX Sequence 1374 BP; 539 A; 358 C; 285 G; 192 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6,39e-63	Length:	1374
Score:	602.00	Matches:	116
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.15%	Mismatches:	0
Query Match:	99.83%	Indels:	0
DB:	12	Gaps:	0

US-09-837-344-20 (1-117) x ADO21919 (1-1374)

```
QY      1 LeuGInGluGInGlnAArgAspLeuGluGlnArgValaAspThrLysLeuAsnLeuGlu 20
         |||||
DB      463 CTCCAGAGCAACACGCGACCTTGAAACAGCGCAAGGCTGACACGAAAAAAACCTGGAA 522
         |||||
QY      21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrgLYArgLeuGluIlePro 40
         |||||
DB      523 CCGCAAAAAGAAACACGCGACCTTCTGCTGAGACCTGTACGGCGCTGAAATCCCA 582
         |||||
QY      41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrgLYIleProHisGlnSerSerLeu 60
         |||||
DB      583 GGTATCGAAGCTCCATCCGAAACGAAACGCGGCTACTACATCCACACAGCAGCAGCTG 642
         |||||
QY      61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80
         |||||
DB      643 CCACAAGATTAATCGCGGAACTCCGCGACAGTAAGAAATCAGCATCATCGAAAAAAC 702
         |||||
QY      81 AsnArgLysSerIleThrThrAsnValaGluGlyArgArgAspIleHisIleGlyHisLeu 100
         |||||
DB      703 AACCCGGAAGACATTACCAACCAAGCTGGAAGGCCCGCCGCACTCACAAAGGCCACTC 762
         |||||
QY      101 GluGluLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 117
         |||||
DB      763 GAAGAAAAGAAAGACGGCTCCATCAACACGAAACGAAAGAACAAAAAGC 813
         |||||
RESULT 5
AA080916
ID      AA080916 standard; cDNA; 1320 BP.
XX
XX      AA080916;
AC
XX      16-OCT-2003 (revised)
DT      25-MAR-2003 (revised)
DT      29-AUG-1995 (first entry)
XX
DE      Plasmodium falciparum LSA-1 repeatless gene cDNA.
XX
XX      Plasmodium falciparum LSA-1 repeatless gene; recombinant poxvirus;
KW      multicomponent multistage malarial vaccines; immunogens;
KW      malaria diagnosis; ss.
XX
XX      Plasmodium falciparum; (pLSARPS.14L.1).
OS
XX      WO9428930-A1.
PN
XX      22-DEC-1994.
XX
XX      10-JUN-1994; 94WO-US006652.
PF
XX      11-JUN-1993; 93US-00075783.
PR      09-JUN-1994; 94US-00257073.
XX
XX      (VIRO-) VIROGENETICS CORP.
```

XX
PI Paoletti E, De Taisne C, Tine JA;
XX WPI; 1995-036113/05.
DR
XX
XX
PT Recombinant poxvirus contg. Plasmodium DNA in non-essential region -
PT useful in vaccines against malaria and for prodn. of Plasmodium
XX immunogens.

XX Claim 3; Fig 11; 183pp; English.

XX
XX AA080916 is the P. falciparum LSA-1 repeatless gene cDNA sequence. New
CC recombinant poxviruses containing either the SERA, ABRA, Plnsp70, AMA-1,
CC Pf25, Pf26, CSP, PfSP2, LSA-1, LSA-1 repeats, MSA-1, MSA-1 (N-
CC terminal p83 or C-terminal gp42) genes, or a combination of these in non-
CC essential regions of their genomes are claimed. These poxviruses (pref.
CC with a virulence reducing genomic deletion or disruption) can be used as
CC vaccines against malaria and for the prodn. of Plasmodium immunogens.
CC These viruses provide multicomponent, multistage vaccines due to their
CC expression of sporozite, liver stage, blood stage and sexual stage
CC proteins. (Updated on 25-MAR-2003 to correct PN field.) (updated on 16-
CC Oct-2003 to standardise OS field)

XX Sequence 1320 BP; 619 A; 128 C; 227 G; 346 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,45e-58	Length:	1320
Score:	564.00	Matches:	110
Percent Similarity:	91.94%	Conservative:	4
Best Local Similarity:	88.71%	Mismatches:	2
Query Match:	93.53%	Indels:	8
DB:	2	Gaps:	1

US-09-837-344-20 (1-117) x AA080916 (1-1320)

```
QY      2 GlnGlnGlnGlnAArgAspLeuGluGln-----ArgLysAla 13
         |||
DB      427 GAAACACAGACAAAGAGATCTTGAAAGAAAAAGCGCGCATCTAAAGAACGAGAGAGCT 486
         |||
QY      14 AspThrLysLysAsnLeuGluArgLysLysGluHisGlyAspIleLeuAlaGluAspLeu 33
         |||||
DB      487 GATCGAAAAAAATTTAGAAAGAAAGAAAGAAACATGAGATGTATTACGAGAGATTAA 546
         |||||
QY      34 TyrGlyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArgGlyTyrgLYr 53
         |||||
DB      547 TATGTCGTTTAGAAATACCGCATATGAACCTTCATCAGAAATGAACGTGATATTAT 606
         |||||
QY      54 IleProHisGlnSerSerLeuProGlnAspAsnArgGlyAsnSerArgAspSerLysGlu 73
         |||||
DB      607 ATACCACTCAATCTTCTTTACTCAGACACACAGAGGAATAGTAGATTCCAAAGAA 666
         |||||
QY      74 IleSerIleIleGluLysThrAsnArgGluSerIleThrThrAsnValaGluGlyArgArg 93
         |||||
DB      667 ATATCTATATATGAAAAAACAATAGAGAAATCTATTCAACAAATGTTGAAGACGAAAG 726
         |||||
QY      94 AspIleHisLysGlyHisLeuGluGluLysLysAspGlySerIleLysProGluGlnLys 113
         |||||
DB      727 GATATACCTTAAGACATCTTGAAGAAAGAAAGAAATGTTCAATTAACAACGAAACAAAA 786
         |||||
QY      114 GluAspLysSer 117
         |||||
DB      787 GAAGATAAATCT 798
         |||||
RESULT 6
ADZ40618
ID      ADZ40618 standard; DNA; 1384 BP.
XX
XX      ADZ40618;
AC
XX      30-JUN-2005 (first entry)
DT
XX      PfCTL/HTL/N multi-epitope construct encoding DNA.
DE
XX
```

KW multi-epitope; immune stimulation; HIV infection; anti-HIV;
XX immunostimulant; vaccine; immunogenicity; gene; ds.
OS Synthetic.
XX
PN WO2005033265-A2.
XX
PD 14-APR-2005.
XX
PF 26-APR-2004; 2004WO-US012732.
XX
PR 25-APR-2003; 2003US-0465229P.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Chesnut RW, Newman MJ, Livingston BD;
XX
DR WPI; 2005-285416/29.
XX
DR P-PSDB; AD240617.
XX
PT New multi-epitope polynucleotide, useful for designing vaccines that
XX provide increased immunogenicity, and for inducing immune responses
PT against HIV in an individual.
XX
PS Disclosure; Fig 181; 261pp; English.
XX
CC The invention relates to a novel multi-epitope polynucleotide. The
XX invention further comprises: a composition comprising the polynucleotide;
CC a cell comprising the polynucleotide; a method for inducing an immune
XX response against human immunodeficiency virus (HIV) in an individual.
CC comprising administering the polynucleotide, composition or cell to the
XX individual; and a method for making the polynucleotide, composition or
CC cell. The multi-epitope polynucleotide has anti-HIV and immunostimulant
XX activities. The polynucleotide is useful for designing vaccines that
CC provide increased immunogenicity, and for inducing an immune response
XX against HIV. This polynucleotide sequence represents an PfCTL/HTL/N multi-
CC -epitope construct encoding DNA of the invention.
XX
SQ Sequence 1384 BP; 351 A; 347 C; 361 G; 325 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0.239 Length: 1384
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: Gaps: 0
US-09-837-344-20 (1-117) x AD240618 (1-1384)
QY 40 ProhlaarlegluLeuPProSerGIuAenGIuArgGlyTYrTYrIleProHISGInSerSer 59
DB 360 CCTGGACCAAGCTCTGCCGTCCGAAACGAAAGAGATCTACTACCTCACCAGAGCAGC 419
QY 60 Leu 60
DB 420 CTC 422
RESULT 7
ADA49438
ID ADA49438 standard; DNA; 1385 BP.
XX
AC ADA49438;
XX
DT 20-NOV-2003 (first entry)
XX
DE Multi-epitope construct #1 DNA.
XX
DS ds; gene; multi-epitope; immunogenic; epitope;
KW major histocompatibility complex; MHC class I; MHC class II;
XX functional epitope.
OS Synthetic.

OS Plasmodium falciparum.
XX
PN US2002119127-A1.
XX
PD 29-AUG-2002.
XX
PF 27-JUN-2001; 2001US-00894018.
XX
PR 28-DEC-1999; 99US-0173390P.
XX
PR 28-DEC-2000; 2000WO-US035568.
XX
PR 16-APR-2001; 2001US-0284221P.
XX
PA (SETT/) SETTE A.
XX
PA (CHES/) CHESNUT R.
XX
PA (LIV/) LIVINGSTON B D.
XX
PA (BAKE/) BAKER D M.
XX
PA (NEWM/) NEWMAN M J.
XX
PA (BROW/) BROWN D H.
XX
PI Sette A, Chesnut R, Livingston BD, Baker DM, Newman MJ, Brown DH;
XX
DR WPI; 2003-615704/58.
XX
DR P-PSDB; ADA49437.
XX
PT Designing multi-epitope construct having major histocompatibility complex
XX class I and II epitope nucleic acids, by selecting mixture of amino acid
XX insertions at junctions of construct to minimize junctional epitopes.
XX
PS Disclosure; Fig 181; 78pp; English.
XX
CC The invention relates to a method of designing multi-epitope constructs
XX comprising major histocompatibility complex (MHC) class I and II (CTL)
XX epitope nucleic acids (CEN), involves sorting CEN, introducing flanking
XX amino acid residue selected from specified amino acid residues given in
XX specification at C+1 position of CEN, introducing amino acid spacer
XX residues between two CEN, and selecting the constructs having less
XX junctional epitopes. The method is useful for designing a multi-epitope
XX construct having multiple epitope nucleic acid. The method avoids or
XX minimises the occurrence of junctional epitopes and maximises the
XX immunogenicity and/or antigenicity of multi-epitope vaccines. The present
XX sequence represents DNA encoding a multi-epitope construct.
XX
SQ Sequence 1385 BP; 351 A; 347 C; 361 G; 326 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0.239 Length: 1385
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: Gaps: 0
US-09-837-344-20 (1-117) x ADA49438 (1-1385)
QY 40 ProhlaarlegluLeuPProSerGIuAenGIuArgGlyTYrTYrIleProHISGInSerSer 59
DB 361 CCTGGACCAAGCTCTGCCGTCCGAAACGAAAGAGATCTACTACCTCACCAGAGCAGC 420
QY 60 Leu 60
DB 421 CTC 423
RESULT 8
ADO24116
ID ADO24116 standard; DNA; 1385 BP.
XX
AC ADO24116;
XX
DT 01-JUL-2004 (first entry)
XX
DE Epigene construct PfCTL/HTL(N) DNA.
XX
DS ds; gene; hepatotropic; virulence; antiinflammatory; anti-HIV; cyostatic;

KW epitope; cytotoxic T lymphocyte; CTL; helper T lymphocyte; HTL;
KW immune response; hepatitis B virus; HBV; HIV; cancer; lymphoma.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT exon 1..1377
FT /*tag= a
XX
PN WO2004031210-A2.
XX
PD 15-APR-2004.
XX
PF 03-OCT-2003; 2003WO-US031303.
XX
PR 03-OCT-2002; 2002US-0415463P.
PR 22-OCT-2002; 2002US-0419773P.
XX
PA (EPRM-) EPIMUNE INC.
PA (GEMV-) GENENCOR INT INC.
PI Sette A, Chenut R, Newman MJ, Livingston BD, Babe LM, Chen Y;
PI DeYoung LM, Huang MTF, Power SD;
XX
DR WPI; 2004-330143/30.
DR P-PSDB; ADO24115.
XX
PT New multi-epitope polynucleotides encoding cytotoxic T lymphocyte and/or
PT helper T lymphocyte epitopes, useful for inducing or stimulating an
PT antiviral or anticancer immune response, especially against hepatitis B
PT virus.
XX
PS Disclosure; SEQ ID NO 266; 401bp; English.
XX
CC The invention relates to a multi-epitope polynucleotide (I) encoding
CC cytotoxic T lymphocyte (CTL) and/or helper T lymphocyte (HTL) epitopes.
CC The composition and methods are useful for inducing or stimulating an
CC immune response against hepatitis B virus (HBV), other viruses (e.g.
CC HIV), or against cancers such as prostate cancer or lymphoma. This
CC sequence corresponds to an example of an epigene construct used in the
CC invention.
XX
SQ Sequence 1385 BP; 351 A; 347 C; 361 G; 326 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.239 Length: 1385
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: Gaps: 0

US-09-837-344-20 (1-117) x ADO24116 (1-1385)

QY 40 ProAlaIleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSer 59
DB 361 CCTGACCAAGCTCTCCGCTCCGAAACGAAAGAGATCTACTACCTCAGAGACGAC 420

QY 60 Leu 60
DB 421 CTC 423

RESULT 9
AA575358/c
ID AA575358 standard; cDNA; 969 BP.
XX
AC AA575358;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #11162.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG11171.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 11162; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 969 BP; 175 A; 188 C; 231 G; 375 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.399 Length: 969
Score: 90.50 Matches: 30
Percent Similarity: 47.66% Conservative: 21
Best Local Similarity: 28.04% Mismatches: 35
Query Match: 15.01% Indels: 21
DB: Gaps: 6

US-09-837-344-20 (1-117) x AA575358 (1-969)

QY 1 LeuGlnGlnGlnGlnArgAspLeuGlnGlnArgTysAlaAspThrLysAsnLeuGlu 20
DB 600 CTCACGCAAGCTCCACAGACCTGCAAGTCCGACTGTAGAGAGAAACTAAC 541

QY 21 ArgLysIlyGlnHisGlyAspIleLeuAlaGlnAspLeuTyrGlyArgLeuGlnIlePro 40
DB 540 AACAGAAAGACATCCACAG-----CAAAACCC 511

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyr-----IleProHisGln 57
DB 510 ATCTGTACGTCAATCATCTCAAGACCAAGAGATATATTCAGAGAGAACTTCCCAACT 451

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (II) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcc_sequences

XX Sequence 969 BP; 375 A; 231 C; 188 G; 175 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.399 Length: 969
 Score: 90.50 Matches: 30
 Percent Similarity: 47.66% Conservative: 21
 Best Local Similarity: 28.04% Mismatches: 35
 Query Match: 15.01% Indels: 21
 DB: 5 Gaps: 6

US-09-837-344-20 (1-117) x AAS92259 (1-969)

QY 1 LeuGInGInGInGInAArgApleuGInAArgLyAAlaAAspThrLybLybAAsnLeuGIn 20
 DB 370 CTCGAGCAAGCTCCAAAGACCTGAGGCTGAGCTGATGTAAGAAAGAACTAAC 429
 QY 21 ArgLybLybGInHIGLyAAspIleuAlaGInAAspLeuTyGlyAArgLeuGInLeuPro 40
 DB 430 AAACGAAAGACATCCACAG-----CAAACCCC 459
 QY 41 AlaIleGInLeuProSerGInAAsnGInAArgLybTyTyT-----IleProHIGIn 57
 DB 460 ATCTGTACTCTACCATCAACAAAGACCAAGATATTATCCAGAGAACTCCCAACT 519
 QY 58 SerSerLeuProGIn---AspAAsnArgLybAAsnSerArgAAspSerLybGInIleSerIle 76
 DB 520 ACCAAGACAGGCCAGCATTCATAATTCAGGAATATACAGAAATGCCAAGATA---CTC 576
 QY 77 IlegLybGInAAsnAArgGInSerIleThrAAsnValGInGlyAArgAAspIleHis 96
 DB 577 CTCGAGAAAGCAAC-----TCCAAAGACACATAATTGTGAG-----ATTGAC 618
 QY 97 LybGInHIGLeuGInLyb 103
 DB 619 CAAGTTGAATGAAGAAAA 639
 RESULT 12
 AAT27052
 ID AAT27052 standard; DNA; 4599 BP.
 XX AAT27052;
 DE 12-OCT-1996 (first entry)
 DT RRP3 telomerase-associated protein gene.
 XX
 DE RRP3 telomerase-associated protein gene.
 XX
 KW Yeast; RRP3 gene; telomerase-associated protein; STR7;
 KW suppressor of telomeric repression-7; telomerase; ribonucleoprotein;
 KW telomere; tumour; pathogen; sperm; reporter gene; drug screening;
 KW antibody; immunosassay; antitumour; antiseptic; contraceptive;
 KW infertility; diagnostic; gene therapy; se.

XX Saccharomyces cerevisiae.
 OS
 XX Key Location/Qualifiers
 FH 835..4092
 FT CDS /*tag=a
 FT /product="RRP3 protein"

XX MO9612811-A2.

XX 02-MAY-1996.

XX 20-OCT-1995; 95WO-US013801.

XX 20-OCT-1994; 94US-00326781.

XX 28-APR-1995; 95US-00431080.

XX (ARCH-) ARCH DEV CORP.

XX Gottschling DE, Singer MS;

XX WPI; 1996-239169/24.

XX P-PSDB; AAR95607.

Novel telomerase associated polypeptide(s) and related nucleic acid -
 useful for detecting e.g. tumour cells or pathogens.

Example 10; Page 287-293; 349pp; English.

The RRP3 gene encodes a telomerase-associated protein (with a fragment,
 CC STR7, shown in AAT27051) found in conjunction with Saccharomyces
 CC cerevisiae telomerase, a ribonucleoprotein required for telomere
 CC replication. Other genes associated with telomerase are given in AAT27045
 CC (telomerase RNA template gene) and AAT27046-50 (other STR genes). These
 CC genes combine with telomerase to repress telomere silencing of gene
 CC expression. Oligonucleotides from the sequence may be used to detect non-
 CC ciliate telomerase-associated genes, e.g. in tumour, pathogen, sperm or
 CC ovum cells. New telomerase-associated sequences may be detected by a
 CC reporter gene expression system linked to an expression-repressing
 CC telomere sequence, and binding compounds, e.g. antibodies, may be
 CC detected by complex formation with telomerase components. The products
 CC may be used as antitumour, antiseptic or contraceptive agents, in
 CC infertility diagnosis, or in gene therapy

Sequence 4599 BP; 1717 A; 779 C; 898 G; 1205 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.04 Length: 4599
 Score: 90.50 Matches: 31
 Percent Similarity: 34.90% Conservative: 21
 Best Local Similarity: 20.81% Mismatches: 54
 Query Match: 15.01% Indels: 43
 DB: 2 Gaps: 4

US-09-837-344-20 (1-117) x AAT27052 (1-4599)

QY 3 GInGInGInAArgAAspLeuGInGInAArgLybAlaAAspThrLybLybAAsnLeuGInAArgLyb 22
 DB 3427 GATTAAGATGAAGAGCTACACAGGAAAAAACTGAAAAAAAGGCACTTAAGAAAAAA 3486
 QY 23 LybGIn----- 24
 DB 3487 AAGAAAGCTTCTTAATAAGAAAGCAAGAAAAAGCAATTAAGAGCAAGAACTATG 3546
 QY 25 -----HIGLybAP 27
 DB 3547 AATTACAATTGGGAATCAATGCGCATGATCATCGTAAACAATTAACAGCCATGCGCAC 3606
 QY 28 Ile---LeuAlaGInAAspLeuTyGlyAArgLeuGInIleProAlaIleGInLeuProSer 46
 DB 3607 ATAATACCGGTACCGATTTTACAACCAATGAATACCCCTGATGAATGATGATCTCTCT 3666
 QY 47 GIn---AAsnGInAArgLybTyTyT---IleProHIGInSerSerLeuProGInAAsnAArg 65

DB 3667 CACGACCTGAAGATCGCTCATTAATACCTCAATTCGATTCCTCCGACACCAAT 3726
||| ||||| ||| |||||
QY 66 GlyAsnSerArgAspSerIleGluLeu-----SerIleIleGluLys 79
||| ||||| ||| |||||
DB 3727 ACAGAGAAAATTCACAAAGAGCTGGTTAGATGAATTCAGATTTTGGGCAAA 3786
||| ||||| ||| |||||
QY 80 ThrAsnArgGluSerIleThrAsnValGluGlyArgArgAspIleHisGlyHis 99
||| ||||| ||| |||||
DB 3787 GATGAAATGACTTACTGCTGTAGGTGATATTACGTTATGATGCACAAAGAGTCAT 3846
||| ||||| ||| |||||
QY 100 LeuGluGluLysLysAspGlySerIle 108
||| ||||| ||| |||||
DB 3847 GTGATCGAAGATGACTGACCCGATATC 3873
||| ||||| ||| |||||
RESULT 13
ADR25659
ID ADR25659 standard; DNA; 4080 BP.
XX
AC ADR25659;
XX
DT 21-OCT-2004 (first entry)
XX
DE Breast cancer prognosis marker #1520.
XX
KW db; breast cancer; prognosis; gene expression; diagnosis.
XX
OS Homo sapiens.
XX
PN MO2004065545-A2.
XX
PD 05-AUG-2004.
XX
PF 15-JAN-2004; 2004MO-US001100.
XX
PR 15-JUN-2003; 2003US-00342887.
XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX
PA (NECA-) NETHERLANDS CANCER INST.
XX
PI Van't Veer LJ, He Y;
XX
DR WPI; 2004-593473/57.
XX
PT Classifying a breast cancer patient according to prognosis comprises
XX
PT determining the similarity between the level of expression of each of
XX
PT five genes in a cell sample taken from patient, to control levels.
XX
PS Disclosure; SEQ ID NO 1520; 226bp; English.
XX
OS The invention relates to a method of classifying a breast cancer patient
XX
CC according to prognosis by determining the similarity between the level of
XX
CC expression of each of five genes for which markers are listed in the
XX
CC specification, in a cell sample taken from the breast cancer patient, to
XX
CC control levels of expression for each respective five genes to obtain a
XX
CC patient similarity value. The methods are useful for classifying a breast
XX
CC cancer patient according to prognosis. Kits and computer program products
XX
CC are useful for data analysis using the diagnostic, prognostic and
XX
CC statistical methods of the invention. This sequence corresponds to a
XX
CC marker used in the method of the invention.
XX
SQ Sequence 4080 BP; 765 A; 1323 C; 1354 G; 638 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.99 Length: 4080
Score: 90.00 Matches: 31
Percent Similarity: 41.67% Conservative: 24
Best Local Similarity: 23.48% Mismatches: 53
Query Match: 14.93% Indels: 24
DB: 13 Gaps: 4
US-09-837-344-20 (1-117) x ADR25659 (1-4080)

QY 2 GlnGluGlnArgAspLeuGluGlnArgLysAlaAspThrLysAsnLeuGluArg 21
||| ||||| ||| |||||
DB 2167 CAGAGAACAGAGGAGAAATGGGGGACAGCGCTGAGTCAGAGAGCAACAGTTCCAGAG 2226
||| ||||| ||| |||||
QY 22 LysLysGluHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGlu--IlePro 40
||| ||||| ||| |||||
DB 2227 GACAAAGAGCCACCCAGAGAGCTGATCGAGACCTCCGAAGACGCTGGACACCTGAC 2286
||| ||||| ||| |||||
QY 41 AlaIleGluLeuProSerGluAsnGluArgGly----- 51
||| ||||| ||| |||||
DB 2287 CTCCTCAAGCTGAGAGCCGAGCAGCGCGGCGCCGACGACGACATGGGCTCCAGAG 2346
||| ||||| ||| |||||
QY 52 TyrTyrIleProHisGlnSerSer-----LeuProGlnAsp 63
||| ||||| ||| |||||
DB 2347 TACCAAGCCCGCCCGGAGACCGAGCTGAGACAGAGTCCGACGCTGAAGCAGAGAC 2406
||| ||||| ||| |||||
QY 64 AsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThrAsnArgLys 83
||| ||||| ||| |||||
DB 2407 AACCGCAACCTGAAGAGCAGAAACAGAGAGCTG-----AACGGGACG 2448
||| ||||| ||| |||||
QY 84 SerIleThrThrAsnValGluGlyArgArgAspIleHisGlyHisLeuGluGluLys 103
||| ||||| ||| |||||
DB 2449 ATCATTAACCTCAGATCCAGAGCGGCCCAAGAGCTTCTCCACAGACCTTCTGAGTCC 2508
||| ||||| ||| |||||
QY 104 LysAspGlySerIleLysProGluGlnLysGluAsp 115
||| ||||| ||| |||||
DB 2509 CTGCTGACAGATCAGCTCCGCTCCCGAGATGAG 2544
||| ||||| ||| |||||
RESULT 14
AAH84231
ID AAH84231 standard; cDNA; 696 BP.
XX
AC AAH84231;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human cell death protective cDNA clone CNI-00717 ORF5, SEQ:193.
XX
KW Cell death protective; apoptosis; necrosis; human; drug screening;
XX
KW cell death-associated disorder; central nervous system disorder;
XX
KW psychiatric disorder; neurological disorder; ischemia-related disorder;
XX
KW stroke; cerebral infarction; ischemic encephalopathy;
XX
KW neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
XX
KW Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
XX
KW vascular disease; opticalmological disorder; diabetic retinopathy;
XX
KW macular degeneration; hypertension; myocardial infarction;
XX
KW atherosclerosis; respiratory disorder; asthma; transgenic animal;
XX
KW chronic obstructive pulmonary disease; neoplastic condition; cancer;
XX
KW benign tumour; anaemia; gastrointestinal disorder; gastritis;
XX
KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
XX
KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;
XX
KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
XX
KW urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS;
XX
KW open reading frame; ORF; ss.
XX
OS Homo sapiens.
XX
PN MO200145638-A2.
XX
PD 28-JUN-2001.
XX
PF 11-DEC-2000; 2000MO-US033547.
XX
PR 14-DEC-1999; 99US-00461697.
XX
PA (COGE-) COGENT NEUROSCIENCE INC.
XX
PI Lo DC, Barney S, Thomas MB, Portbury SD, Putanam K, Katz LC;
XX
DR WPI; 2001-390297/41.
XX
DR P-PSDB; AAG98701.
XX
PT Novel protective sequence polynucleotides and polypeptides, used to

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 07:30:56 ; Search time 92.9515 Seconds

(Without alignments)
2237.454 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603
Sequence: 1 LQEQRDLEQRKADTKKLE.....GHLEKKDSIKPEQKEDKS 117

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Egapop 10.0 , Egapext 0.5	
Dgapop 6.0 , Dgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool.h/US09837344/runat.05122005.133533.763/app.query.fasta.1.718
-DB=Issued_Patents_NA -QFMT=fasta -SUFFIX=rni -MINMATCH=0.1 -LOOFCF=0
-LOOPEXT=0 -INIT=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09837344 @CGN 1.1 237 @runat.05122005.133533.763 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOB=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/prodata/1/lna/1.COMB.seq.*
2: /cgn2_6/prodata/1/lna/5.COMB.seq.*
3: /cgn2_6/prodata/1/lna/6A.COMB.seq.*
4: /cgn2_6/prodata/1/lna/6B.COMB.seq.*
5: /cgn2_6/prodata/1/lna/H.COMB.seq.*
6: /cgn2_6/prodata/1/lna/PCTUS.COMB.seq.*
7: /cgn2_6/prodata/1/lna/PP.COMB.seq.*
8: /cgn2_6/prodata/1/lna/RE.COMB.seq.*
9: /cgn2_6/prodata/1/lna/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	603	100.0	950	3	US-08-098-327E-32 Sequence 32, App1
2	603	100.0	950	3	US-08-462-625-32 Sequence 32, App1
3	603	100.0	1482	3	US-08-098-327E-41 Sequence 41, App1
4	603	100.0	1482	3	US-08-462-625-41 Sequence 41, App1
5	603	100.0	1493	3	US-08-098-327E-38 Sequence 38, App1
6	603	100.0	1493	3	US-08-462-625-38 Sequence 38, App1
7	600	99.5	1482	3	US-08-098-327E-45 Sequence 45, App1
8	600	99.5	1482	3	US-08-462-625-45 Sequence 45, App1
9	564	93.5	1320	2	US-08-257-073-15 Sequence 15, App1

10	111	18.4	71	2	US-08-257-073-130	Sequence 130, App
11	97	16.1	2130	3	US-09-248-796A-4942	Sequence 4942, Ap
12	93	15.4	73519	3	US-09-949-016-16344	Sequence 16344, A
13	93	15.4	105919	3	US-09-949-016-11769	Sequence 11769, A
14	90.5	15.0	4599	2	US-08-431-080-27	Sequence 27, App1
15	90.5	15.0	4599	2	US-08-938-534-27	Sequence 27, App1
16	90.5	15.0	4599	2	US-09-345-294-27	Sequence 27, App1
17	89.5	14.8	696	3	US-09-461-697-193	Sequence 193, App
18	89.5	14.8	699	3	US-09-461-697-193	Sequence 193, App
19	89.5	14.8	717	3	US-09-461-697-189	Sequence 189, App
20	89.5	14.8	717	3	US-09-461-697-187	Sequence 187, App
21	89.5	14.8	819	3	US-09-461-697-185	Sequence 185, App
22	89.5	14.8	1669	3	US-09-461-697-184	Sequence 184, App
23	86	14.3	1344	3	US-09-710-279-2899	Sequence 2899, Ap
24	86	14.3	1353	3	US-09-710-279-3235	Sequence 3235, Ap
25	86	14.3	3001	3	US-09-710-279-3813	Sequence 3813, Ap
26	86	14.3	3974	3	US-09-710-279-4245	Sequence 4245, Ap
27	85.5	14.2	3739	3	US-09-882-274-1	Sequence 1, App1
28	85	14.1	19728	3	US-09-949-016-12506	Sequence 12506, A
29	85	14.1	19728	3	US-09-949-016-15767	Sequence 15767, A
30	84.5	14.0	1785	3	US-09-949-016-2716	Sequence 2716, Ap
31	84.5	14.0	1758	3	US-09-949-016-756	Sequence 756, App
32	84.5	14.0	3645	2	US-08-663-112-1	Sequence 1, App1
33	84	13.9	5874	3	US-09-561-709B-2	Sequence 2, App1
34	83.5	13.8	1685	3	US-10-104-047-1957	Sequence 1957, Ap
35	83.5	13.8	73519	3	US-09-949-016-16344	Sequence 16344, A
36	83.5	13.8	105919	3	US-09-949-016-11769	Sequence 11769, A
37	82.5	13.7	2471	3	US-09-915-060A-3	Sequence 3, App1
38	81.5	13.5	5788	3	US-09-949-016-12498	Sequence 12498, A
39	81.5	13.5	5788	3	US-09-949-016-14458	Sequence 14458, A
40	81.5	13.5	5870	3	US-09-949-016-15247	Sequence 15247, A
41	81	13.4	4248	3	US-10-164-595-53	Sequence 53, App1
42	81	13.4	28806	3	US-09-949-016-13217	Sequence 13217, A
43	80.5	13.3	431	3	US-09-513-999C-1630	Sequence 1630, App
44	80.5	13.3	1353	3	US-08-956-171E-521	Sequence 521, App
45	80.5	13.3	1353	3	US-08-781-986A-521	Sequence 521, App

ALIGNMENTS

RESULT 1
US-08-098-327E-32
; Sequence 32, Application US/08098327E
; Patent No. 6270771
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; TITLE OF INVENTION: DRIUHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; City: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,327E
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-32

Alignment Scores:
Pred. No.: 4,966-66 Length: 950
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-08-098-327E-32 (1-950)

QY 1 LeungIngluInglInarAspleuGluInarGlybAlaAspThrLysLysAsnLeuGlu 20
DB 600 TTGCAGAGACGACCAAGAGATTAGAACAAAGAGAGCTGATACGAAAAAATTTAGAA 659

QY 21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyGlyArgLeuGluLeuPro 40
DB 660 AGAAAAAGAAACATGAGATATATTAGCAGAGATTATATATGTCGTTTAAATAATACCA 719

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyTyTyrIleProHisGlnSerSerLeu 60
DB 720 GCTATAGAACTTCATCAGAAAAATGACGTGATATTATATCCACATCAATCTCTTTA 779

QY 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80
DB 780 CCTCAGACACACAGAGGAGATAGTAGAGATTCCAGAAATATCTATATAGAAAAACA 839

QY 81 AsnArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHisLysGlyHisLeu 100
DB 840 AATAGAGATCTATTACCAAAATGTTGAAGACGAGAGGATATACATTAAGACATCTT 899

QY 101 GluGluLysLysAspGlySerIleLysProGluGluLysGluAspLysSer 117
DB 900 GAAGAAAAAGAAAGATGTTCAATTAACACAGAACAAAAAGAGATTAATCT 950

RESULT 2
US-08-462-625-32
Sequence 32, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462, 625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098, 327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-32

Alignment Scores:
Pred. No.: 4,966-66 Length: 950
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-08-462-625-32 (1-950)

QY 1 LeungIngluInglInarAspleuGluInarGlybAlaAspThrLysLysAsnLeuGlu 20
DB 600 TTGCAGAGACGACCAAGAGATTAGAACAAAGAGAGCTGATACGAAAAAATTTAGAA 659

QY 21 ArgLysLysGluHisGlyAspIleLeuAlaGluAspLeuTyGlyArgLeuGluLeuPro 40
DB 660 AGAAAAAGAAACATGAGATATATTAGCAGAGATTATATATGTCGTTTAAATAATACCA 719

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyTyTyrIleProHisGlnSerSerLeu 60
DB 720 GCTATAGAACTTCATCAGAAAAATGACGTGATATTATATCCACATCAATCTCTTTA 779

QY 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80
DB 780 CCTCAGACACACAGAGGAGATAGTAGAGATTCCAGAAATATCTATATAGAAAAACA 839

QY 81 AsnArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHisLysGlyHisLeu 100
DB 840 AATAGAGATCTATTACCAAAATGTTGAAGACGAGAGGATATACATTAAGACATCTT 899

QY 101 GluGluLysLysAspGlySerIleLysProGluGluLysGluAspLysSer 117
DB 900 GAAGAAAAAGAAAGATGTTCAATTAACACAGAACAAAAAGAGATTAATCT 950

RESULT 3
US-08-098-327E-41
Sequence 41, Application US/08098327E
Patent No. 6270771
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-098-327E-41
Alignment Scores:
Pred. No.: 9.13e-66 Length: 1482
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 3
DB: 3
US-09-837-344-20 (1-117) x US-08-098-327E-41 (1-1482)
QY 1 LeuGInGluGlnGlnAArgApleuGluGlnAArgApleuAAspThrIlySlySAsnLeuGlu 20
DB 610 TTGCAAGAGCAGCAAGAGATTGAGACAAAGAGAGGCTGATCGAAAAAATTTTGA 669
QY 21 ArgLySlySgluHIsGlyAspIleLeuAlaGluAspLeuTyrgIyArgLeuGluIlePro 40
DB 670 AGAAAAAGCAATGAGATATATTAGCAGAGATTATATGTTGTTAGAAATGCCA 729
QY 41 AlaIleGluLeuProSerGluAsnGluArgIlyTyrrIleProHIsGlnSerSerLeu 60
DB 730 GCTATGAACTTCATCAGAAAAATGAACGTGATATTATACACATCAATCTTCTTA 789
QY 61 ProGluAspAsnArgIyAsnSerArgAspSerIlySgluIleSerIleIleGluIlyThr 80
DB 790 CCTCAGGACACAGAGGAAATAGTAGATTCCAGGAAATATCTATATAGAAAAACA 849
QY 81 AsnArgIlySerIleThrAsnValGluGlyArgAAspIleHIsIlySgluHIsLeu 100
DB 850 AATAGAGAACTTATTAACAATAATTGTTGAAGAGCAAGAGATATACATAAGGACATCTT 909
QY 101 GluGluIlySlySAspGlySerIleIlySProGluGluIlySgluAspIlySer 117

DB 910 GAAGAAAAAGATGCTTCATTAATAACCAAGACAAAAAGAGATTAATCT 960
RESULT 4
US-08-462-625-41
Sequence 41, Application US/08462625
Patent No. 6319502
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRULHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/098,327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
US-08-462-625-41
Alignment Scores:
Pred. No.: 9.13e-66 Length: 1482
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 0
US-09-837-344-20 (1-117) x US-08-462-625-41 (1-1482)
QY 1 LeuGInGluGlnGlnAArgApleuGluGlnAArgApleuAAspThrIlySlySAsnLeuGlu 20
DB 610 TTGCAAGAGCAGCAAGAGATTGAGACAAAGAGAGGCTGATCGAAAAAATTTTGA 669
QY 21 ArgLySlySgluHIsGlyAspIleLeuAlaGluAspLeuTyrgIyArgLeuGluIlePro 40

Db 670 AGAAAAAGAAACATGAGATATATTACAGAGATTATATGCTGTTTGAATATCCA 729
Qy 41 AlallegluleupProsergiuaenGluArgLYTYrTyrilleProHisgInSerSerleu 60
Db 730 GCTTAAACCTTCATCAGAAAATGAAAGTGTATATATATACCAATCATCTTCTTTA 789
Qy 61 ProGlnaPasnArgLYaenSerArgAspSerlysgLulleSerilleleGluLysThr 80
Db 790 CCTCAGGACAAACAGAGGAATAGTAGATTCCAAAGAAATATCTATATATGAAAAACA 849
Qy 81 AsnArgGluSerilleThrThrAsnValGluGlyArgArgAspIleHislysgLYHisleu 100
Db 850 AATGAGAAATCTATTACAAACAAATGTTGAAGACAGAAAGGATATATCAATAAGACATCTT 909
Qy 101 GluGluLysLYsaAspGlySerilleYspProGluGlnLYsaAspLYsSer 117
Db 910 GAAGAAAAGAAACATGCTTCAATTAACCAAGAAACAAAAGATTAATCT 960

RESULT 5
US-08-098-327E-38
Sequence 38, Application US/08098327E
Patent No. 6270771

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098.327E
FILING DATE: 24-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
PUBLICATION NUMBER: WO 92/13884
DOCUMENT NUMBER: 20-AUG-1992
US-08-098-327E-38

Alignment Scores:
Pred. No.: 9,23e-66 Length: 1493
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-08-098-327E-38 (1-1493)

Qy 1 LeuGlnGluGlnGlnArgAspLeuGluGlnArgLYsaAspThrLYsLYsaAsnLeuGlu 20
Db 607 TTCCAGAGCAGCAAGAAAGATTTAGAAACAAGAAAGGCTGATACGAAAAAATTTTGAA 666
Qy 21 ArgLYsLYsgLunHisgLYaAspIleleuAlleGluAspLeuYrGlyArgLeuGluIlePro 40
Db 667 AGAAAAAGAAACATGAGATATATTAGCAGAGATTTATATGCTGTTTGAATATCCA 726
Qy 41 AlallegluleupProsergiuaenGluArgLYTYrTyrilleProHisgInSerSerleu 60
Db 727 GCTATAGAACTTCATCAGAAAATGAAACGTGTATATATATACCAATCATCTTCTTTA 786
Qy 61 ProGlnaPasnArgLYaenSerArgAspSerlysgLulleSerilleleGluLysThr 80
Db 787 CCTCAGGACAAACAGAGGAATAGTAGATTCCAAAGAAATATCTATATATGAAAAACA 846
Qy 81 AsnArgGluSerilleThrThrAsnValGluGlyArgArgAspIleHislysgLYHisleu 100
Db 847 AATGAGAAATCTATTACAAACAAATGTTGAAGACAGAAAGGATATATCAATAAGACATCTT 906
Qy 101 GluGluLysLYsaAspGlySerilleYspProGluGlnLYsaAspLYsSer 117
Db 907 GAAGAAAAGAAACATGCTTCAATTAACCAAGAAACAAAAGATTAATCT 957

RESULT 6
US-08-462-625-38
Sequence 38, Application US/08462625
Patent No. 6319502

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine
APPLICANT: DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
TITLE OF INVENTION: OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462.625
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/098.327
FILING DATE: 24-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid

Db 607 ATACACATCATCTTCTTACTCAGACACAGAGGAATGATGATTCACAGAA 666
Qy 74 IISerIleIleGluIyThrAsnArgGluSerIleThrThrAsnValGluGlyArgArg 93
Db 667 ATATCTATATAGAAAAAACAATAGAAATCTATTACACAAATGTTGAAGACGACAG 726
Qy 94 ApplIleIleGluIyHISleuGluGluIyLysAspGlySerIleLysProGluGlnIly 113
Db 727 GATATACATTAAGACACTTGTGAAGAAAGAAAGATGTTCAATAAAAACCAAGAA 786
Qy 114 GluAspLysSer 117
Db 787 GAAGATTAATCT 798
RESULT 10
US-08-257-073-130
Sequence 130, Application US/08257073
Patent No. 5766597
GENERAL INFORMATION:
APPLICANT: Paolucci, Enzo
APPLICANT: de Taisne, Charles
TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue, 25th Floor
CITY: New York
STATE: New York
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/257,073
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/852,305
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/672,183
FILING DATE: 20-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2570
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 130:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-257-073-130
Alignment Scores:
Pred. No.: 4.87e-06 Length: 71
Score: 111.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.41% Indels: 0
Gaps: 0

US-09-837-344-20 (1-117) x US-08-257-073-130 (1-71)
Qy 69 ArgAspSerIleGluIleSerIleIleGluIyThrAsnArgGluSerIleThrThrAsn 88
Db 1 AGAGATTCCAAAGAAATATCTATTAATGAAAAAACAATAGAAATCTATTACAACAAT 60
Qy 89 ValGluGly 91
Db 61 GTTGAAAGGA 69
RESULT 11
US-09-248-796A-4942
Sequence 4942, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 4942
LENGTH: 2130
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-4942
Alignment Scores:
Pred. No.: 0.0294 Length: 2130
Score: 97.00 Matches: 37
Percent Similarity: 44.00% Conservative: 29
Best Local Similarity: 24.67% Mismatches: 36
Query Match: 16.09% Indels: 48
Gaps: 9
US-09-837-344-20 (1-117) x US-09-248-796A-4942 (1-2130)
Qy 11 ArgLysAlaIleAspThrIySlySAsnLeuGluArgLysGluHISleGlyAspIle----- 28
Db 1384 AAGAAAGTAGATCTGATCTTAATATTCAGAAAGAGAAATTTATCCAAATTAAGAG 1443
Qy 29 -----LeuAlaGluAspLeuTyrgLyArgLeuGluIleProAla 41
Db 1444 CCACTAAAAGATGAAAAACAATTAAGAAGACTTGCTC---AAAGTTGAACCCATCC 1500
Qy 42 -----IleGluLeuProSerGluAsnGluArgIyTyrrTyrIle 54
Db 1501 AAGATGAGACCGCGGAGTTGAAGTTCCATCCAAAGATGAACACCATTCAGATTC 1560
Qy 55 ProHISleIleSerLeuProGlnAspAsnArgGlyAsnSerArgAspSerIys----- 72
Db 1561 CAAATATTTGATAGT---ACCACTGATAT---GGCAAACTTAAAGATGAGAGATTAAT 1614
Qy 73 -----GluIleSerIleIleGluIySlyThrAsnArgGluSer 84
Db 1615 GTTCTTAAGTTGATTTCAAGTTGATTAAGTATGTTACCAAGACATCAAGATGCT 1674
Qy 85 Ile-----ThrThrAsnValGluGlyArgAsp 94
Db 1675 ATTAAAAAGTAGAAGAAAGCTGTGTAAGAAAGACGACGACGAAAGGTAAAGAAAGAC 1734
Qy 95 -----IleHISlyGlyHISleuGluGluIySlyAspGlySerIleIly 109
Db 1735 AATGAAAAAGTTACTGTTCCAAAGACAAATGAGAGAAATTAAGATTAAGCGTTCAT 1794
Qy 110 ProGlu-----GlnLysGluAsp 115
Db 1795 GAAGCCAAAGATTAAGTTCAAGAAAGATGAT 1824

```
RESULT 12
US-09-949-016-16344
; Sequence 16344, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16344
; LENGTH: 73519
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(73519)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16344

Alignment Scores:
Pred. No.: 12          Length: 73519
Score: 93.00          Matches: 34
Percent Similarity: 45.22%      Conservative: 18
Best Local Similarity: 29.57%    Mismatches: 49
Query Match: 15.42%            Indels: 14
DB: 3                      Gaps: 4

US-09-837-344-20 (1-117) x US-09-949-016-16344 (1-73519)
QY 2 GINGUGLGLNARGAERPLEUGLUGLNAATGYSALAERPLHLYEYASNLGUAATG 21
DB 44860 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 44913
QY 22 LYSYGLNHLGIAERPLLEUWLAGIUAERPLEUTYRGLYARGLEUGLUILEPROLA 41
DB 44914 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 44970
QY 42 ILEGULEUPROSERGLUANGIUAATGTYRTYRTILEPROHISGLINSEERLEUPRO 61
DB 44971 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 45000
QY 62 GINAERPAENARGGLYANSEERARGAERSELYSGLIULESERILETLEGLUYSTRAN 81
DB 45001 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 45060
QY 82 ARGGLUSERILETHTRASNVALIGLUGIYARGARGAPRIEHLIELYSGLYHISLEUGI 101
DB 45061 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 45117
QY 102 GIULYLYEAPRGLYSEERILEYSPROGLUGLNLUGIUAERPLYE 116
DB 45118 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 45162

RESULT 13
US-09-949-016-11769
; Sequence 11769, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11769
; LENGTH: 105919
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(105919)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11769

Alignment Scores:
Pred. No.: 19.9        Length: 105919
Score: 93.00          Matches: 34
Percent Similarity: 45.22%      Conservative: 18
Best Local Similarity: 29.57%    Mismatches: 49
Query Match: 15.42%            Indels: 14
DB: 3                      Gaps: 4

US-09-837-344-20 (1-117) x US-09-949-016-11769 (1-105919)
QY 2 GINGUGLGLNARGAERPLEUGLUGLNAATGYSALAERPLHLYEYASNLGUAATG 21
DB 77261 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77314
QY 22 LYSYGLNHLGIAERPLLEUWLAGIUAERPLEUTYRGLYARGLEUGLUILEPROLA 41
DB 77315 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77371
QY 42 ILEGULEUPROSERGLUANGIUAATGTYRTYRTILEPROHISGLINSEERLEUPRO 61
DB 77372 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77401
QY 62 GINAERPAENARGGLYANSEERARGAERSELYSGLIULESERILETLEGLUYSTRAN 81
DB 77402 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77461
QY 82 ARGGLUSERILETHTRASNVALIGLUGIYARGARGAPRIEHLIELYSGLYHISLEUGI 101
DB 77462 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77518
QY 102 GIULYLYEAPRGLYSEERILEYSPROGLUGLNLUGIUAERPLYE 116
DB 77519 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77563

RESULT 14
US-08-431-080-27
; Sequence 27, Application US/08431080
; Patent No. 5698686
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; APPLICANT: Singer, Miriam S.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

      OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
      SOFTWARE: PatentIn Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/431,080
        FILING DATE: Concurrently Herewith
        CLASSIFICATION: 514
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: SN 08/326,781
          FILING DATE: October 20, 1994
        CLASSIFICATION: 514
        ATTORNEY/AGENT INFORMATION:
          NAME: Parker, David L.
          REGISTRATION NUMBER: 32,165
          REFERENCE/DOCKET NUMBER: AC/D-155/PAR
          TELECOMMUNICATION INFORMATION:
            TELEPHONE: (512) 418-3000
            TELEFAX: (713) 789-2679
            TELEX: 79-0924
        INFORMATION FOR SEQ ID NO: 27:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 4599 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: single
            TOPOLOGY: linear
      US-08-431-080-27

Alignment Scores:
Pred. No.:           0.55           Length:         4599
Score:              90.50           Matches:        31
Percent Similarity: 34.90%          Conservative:   21
Best Local Similarity: 20.81%       Mismatches:    54
Query Match:        15.01%          Indels:        43
DB:                  2             Gaps:          4

US-09-837-344-20 (1-117) x US-08-431-080-27 (1-4599)

Oy      3  GLUGINGLNAAGBAPLEUGLUGLNGLYAIAEPHTIYELYBAENLEUGLUARGYS  22
Db      3427  GATAGATGAAAGAGCTTCAACGGMAAACCTGAAAAAAGCGACTTAAGAAAAA  3486
Oy      23  LYEGHU-----  24
Db      3487  AAGAAGCTTTCTTAAATAAGAAAGCAAAGACAACAAAAGCAATTAAAGSAGCAAACTATG  3546
Oy      25  -----HIGLYABP  27
Db      3547  AATTTCACATTGGGAATCATATGCCATGATCATCGTAAACAATTAACAGCCATAGCGAC  3606
Oy      28  ILE---LEUALAGLUAPLEUTYCYLYRGYLEUGULIEPROHAIILEGLULEUPROSER  46
Db      3607  ATTAATACCGGTACCGATTTTACACAACAATGAMAATACCCTTGATGAATGAAGACTTCCCTC  3666
Oy      47  GLU---AENGILVARGLYTYR-TYRIEPRHISGINSERSEAEUFGGLASPBANARG  65
Db      3667  CACGCACCTGGAAGATGCGCTCATTAATCTCATTAATTTGATCTTGCCGTGACACGAACT  3726
Oy      66  GLYASNSERAYASPSESLYEGLUILE-----SERIEILEGLULYS  79
Db      3727  ACAAGGAAAAATTCAACAAAAGTGTGGTTTAGTGAATTCATGAGATTTTGGCGAAA  3786
Oy      80  THRASNATGJUSERIETHTHRASNAVILGIUGIYARGABPRLIEHLYSGLYHI  99
Db      3787  GATGAAAAATGACTTCTGCTCTGTAGTGATATTAAACGGTTATGATGACACAAGAAGTCAT  3846
Oy      100  LEUGLUGLUYLSLYEABPGIYSEIRLE  108
Db      3847  GTGATCGAAGATACCTACGCGCGATATC  3873

RESULT 15
US-08-938-534-27
; Sequence 27, Application US/08938534
; Patent No. 5916752
; GENERAL INFORMATION:

```

```

1  APPLICANT: Gottschling, Daniel E.
2  APPLICANT: Singer, Miriam S.
3  TITLE OF INVENTION: Telomerase Compositions and Methods
4  NUMBER OF SEQUENCES: 32
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE: Arnold, White & Durkee
7  STREET: P.O. Box 4433
8  CITY: Houston
9  STATE: TEXAS
10 COUNTRY: UNITED STATES OF AMERICA
11 ZIP: 77210
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: FLOPPY disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
17 SOFTWARE: PatentIn Release #1.0, Version #1.30
18
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/938,534
21 FILING DATE: 26-SEP-1997
22 CLASSIFICATION: 536
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: 08/431,080
26
27 FILING DATE:
28 APPLICATION NUMBER: SN 08/326,781
29 FILING DATE: October 20, 1994
30 ATTORNEY/AGENT INFORMATION:
31 NAME: Parker, David L.
32 REGISTRATION NUMBER: 32,165
33 REFERENCE/DOCKET NUMBER: ARCD:155/PAR
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: (512) 418-3000
36 TELEFAX: (713) 789-2679
37
38 TELEX: 79-0924
39
40 INFORMATION FOR SEQ ID NO: 27:
41
42 SEQUENCE CHARACTERISTICS:
43 LENGTH: 4599 base pairs
44 TYPE: nucleic acid
45 STRANDEDNESS: single
46 TOPOLOGY: linear
47
48 US-08-938-534-27
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
```

```

Db      3727 ACAAGGAAAAATTCACAAAAAGCTGTGTTAGATGAATTCAAGATTGCGCAAA 3786
Qy      80  ThrAsnArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHisGlyHis 99
          :::      ::::      |||      ::|||
Db      3787 GATGAAAATGACTTACTGTCTGTAGTGATATTAAAGGTTATGATGACAAGAAGGTCA 3846
Qy      100 LeuGluGluLysLysAspGlySerIle 108
          :::      |||      |||
Db      3847 GTGATCGAAGATACTGACGCCGATATC 3873

```

Search completed: December 7, 2005, 12:45:14
 Job time : 127.951 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2005, 10:33:33 ; Search time 409.905 Seconds
(without alignments)
2360.343 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603

Scoring table:

BL0SUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgnt2_1/USPTO.spool.b/US09837344/runat.05122005.133535.839/app.query.fasta_1.718
-DB=Published_Application.NA.Main -QWTF=Fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=dlbsum62
-TRANS=numa40.cdi -LIST=45 -DCCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09837344@CGN 1.1.1364 @runat.05122005.133535.839
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOC -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Application NA.Main:

1: /cgnt2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgnt2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
3: /cgnt2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
4: /cgnt2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
5: /cgnt2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
6: /cgnt2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgnt2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgnt2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgnt2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
10: /cgnt2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	603	100.0	US-09-837-344-32	Sequence 32, Appl
2	603	100.0	US-09-900-963-32	Sequence 32, Appl
3	603	100.0	US-09-837-344-41	Sequence 41, Appl
4	603	100.0	US-09-900-963-41	Sequence 41, Appl
5	603	100.0	US-09-837-344-38	Sequence 38, Appl
6	603	100.0	US-09-900-963-38	Sequence 38, Appl
7	600	99.5	US-09-837-344-45	Sequence 45, Appl
8	600	99.5	US-09-900-963-45	Sequence 45, Appl

9	94	15.6	1385	3	US-09-894-018-122	Sequence 122, App
10	94	15.6	1385	8	US-10-474-960A-122	Sequence 122, App
11	90.5	15.0	969	9	US-10-450-763-7484	Sequence 7484, App
12	90.5	15.0	969	9	US-10-450-763-11162	Sequence 11162, A
13	90.5	15.0	969	9	US-10-450-763-28063	Sequence 28063, A
14	90	14.9	4080	6	US-10-172-118-1520	Sequence 1520, App
15	90	14.9	4080	7	US-10-342-887-1520	Sequence 1520, App
16	90	14.9	4080	9	US-10-756-149-1669	Sequence 1669, App
17	89.5	14.8	666	3	US-09-922-261-193	Sequence 193, App
18	89.5	14.8	699	3	US-09-922-261-191	Sequence 191, App
19	89.5	14.8	717	3	US-09-922-261-189	Sequence 189, App
20	89.5	14.8	714	3	US-09-922-261-187	Sequence 187, App
21	89.5	14.8	719	3	US-09-922-261-185	Sequence 185, App
22	89.5	14.8	1669	3	US-09-922-261-184	Sequence 184, App
23	87.5	14.5	700	5	US-10-040-739-1385	Sequence 1385, App
24	87	14.4	903	6	US-10-369-493-29682	Sequence 29682, A
25	86.5	14.3	601	7	US-10-021-323-563	Sequence 563, App
26	86.5	14.3	3684	10	US-11-097-143-32615	Sequence 32615, A
27	85.5	14.2	2298	7	US-10-052-482-186	Sequence 186, App
28	85.5	14.2	3477	3	US-09-969-347-221	Sequence 221, App
29	85.5	14.2	3477	9	US-10-843-641A-8350	Sequence 8350, App
30	85.5	14.2	3739	3	US-09-882-274-1	Sequence 1, Appl1
31	85.5	14.2	3740	7	US-10-052-482-185	Sequence 185, App
32	85.5	14.2	3740	8	US-10-484-577-680	Sequence 680, App
33	85.5	14.2	3756	6	US-10-133-013-255	Sequence 255, App
34	85	14.1	613	7	US-10-424-599-28673	Sequence 28673, A
35	85	14.1	11065	8	US-10-723-860-2002	Sequence 2002, App
36	85	14.1	11065	9	US-10-756-149-1888	Sequence 1888, App
37	84.5	14.0	729	9	US-10-450-763-2486	Sequence 2486, App
38	84.5	14.0	1758	9	US-10-450-763-11623	Sequence 11623, A
39	84.5	14.0	2108	3	US-09-962-832-225	Sequence 225, App
40	84.5	14.0	2108	3	US-10-240-425-1294	Sequence 1294, App
41	84.5	14.0	2108	9	US-10-843-641A-6111	Sequence 6111, App
42	84.5	14.0	2108	9	US-09-292-758-85	Sequence 85, Appl
43	84.5	14.0	3645	6	US-10-388-410-4	Sequence 4, Appl1
44	84.5	14.0	630	4	US-09-925-065A-788059	Sequence 788059, A

ALIGNMENTS

RESULT 1
US-09-837-344-32
; Sequence 32, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
; APPLICANT: GUERIN-MARCHAND, Claudine
; DRUTHE, Pierre
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
; HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CARABABLE
; OF STIMULATING THE T LYMPHOCYTES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/837,344
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/462,625
; FILING DATE: <Unknown>
; APPLICATION NUMBER: FR 91 01286
; FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-837-344-32

Alignment Scores:
Pred. No.: 4,27e-66 Length: 950
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-837-344-32 (1-950)

QY 1 LeuGInGluGInGlnArgAspLeuGluGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
DB 600 TTGCAAGACGACGAAAGAGATTAGAACAAAGAGAGCGTGATACGAAAAAATTTAGAA 659

QY 21 ArgLysLysGluHisGlyAspLleLeuAlaGluAspLeuTyrgLysArgLeuGluLeuPro 40
DB 660 AGAAAAAGGACATGAGATATATTAGCAGAGAGATTATATGTCGTTTGAATAACCA 719

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrgLysIleProHisGlnSerSerLeu 60
DB 720 GCATATGAACCTTCATCAGAAATGAACTGATATTATATACCATCATCATCTCTTTA 779

QY 61 ProGlnAspAsnArgLysAsnSerArgAspSerLysGluIleSerLleIleGluLysThr 80
DB 780 CCTCAGACACACAGAGGAAATAGTAGATTCCAGAGAAATATCTATATAGAAAAACA 839

QY 81 AsnArgLysSerLleThrThraenValGluGlyArgArgAspLleHisLysGlyHisLeu 100
DB 840 AATAGAGATCTATTCAACAAATGTTGAAGACGAAAGGATATACATTAAGACATCTT 899

QY 101 GluGluLysLysAspGlySerLleLysProGluGluLysGluAspLysSer 117
DB 900 GAAGAAAGAAAGATGTTCAATMAAACGAAACAAAGAGATMAATCT 950

RESULT 2
US-09-900-963-32
Sequence 32, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burne, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-900-963-32

Alignment Scores:
Pred. No.: 4,27e-66 Length: 950
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-900-963-32 (1-950)

QY 1 LeuGInGluGInGlnArgAspLeuGluGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
DB 600 TTGCAAGACGACGAAAGAGATTAGAACAAAGAGAGCGTGATACGAAAAAATTTAGAA 659

QY 21 ArgLysLysGluHisGlyAspLleLeuAlaGluAspLeuTyrgLysArgLeuGluLeuPro 40
DB 660 AGAAAAAGGACATGAGATATATTAGCAGAGAGATTATATGTCGTTTGAATAACCA 719

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrgLysIleProHisGlnSerSerLeu 60
DB 720 GCATATGAACCTTCATCAGAAATGAACTGATATTATATACCATCATCATCTCTTTA 779

QY 61 ProGlnAspAsnArgLysAsnSerArgAspSerLysGluIleSerLleIleGluLysThr 80
DB 780 CCTCAGACACACAGAGGAAATAGTAGATTCCAGAGAAATATCTATATAGAAAAACA 839

QY 81 AsnArgLysSerLleThrThraenValGluGlyArgArgAspLleHisLysGlyHisLeu 100
DB 840 AATAGAGATCTATTATCAACAAATGTTGAAGACGAAAGGATATACATTAAGACATCTT 899

QY 101 GluGluLysLysAspGlySerLleLysProGluGluLysGluAspLysSer 117
DB 900 GAAGAAAGAAAGATGTTCAATMAAACGAAACAAAGAGATMAATCT 950

RESULT 3
US-09-837-344-41
Sequence 41, Application US/09837344
Patent No. US20020041882A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUIHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE

HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 41:

US-09-837-344-41

Alignment Scores:
Pred. No.: 7.61e-66 Length: 1482
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-837-344-41 (1-1482)

Qy 1 LeuGInGluGInGInaGAspLeuGluGInaGlyAspIleThrlLysLysAsnLeuGlu 20
Db 610 TTGCAAGACGACGAAAGATTGAGACAAAGAGGCTGTATGCAAAAAAATTTGAAA 669

Qy 21 ArgLysLysGluHileGlyAspIleLeuAlaGluAspLeuTyrgLysArgLeuGluIlePro 40
Db 670 AGAAAAAGAAATCGATGATATATTGACAGAGATTATATGTCGTTTGAATACCA 729

Qy 41 AlaileGluLeuProSerGluAsnGluArgGlyTyrrLysIleProHileGlnSerLeu 60
Db 730 GCTATAGAACTTCATCAGAAAAATGAACTGATATTATATACACATCAATCTTCTTA 789

Qy 61 ProGlnAspAsnAspGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80
Db 790 CCTCAGGACACAGAGGSAATGATGAGATTCAGAGAAATATCTTATATGAAAAACA 849

Qy 81 AsnArgLysSerIleThrlrAsnValGluGlyArgArgAspIleHisLysGlyHisLeu 100
Db 850 AATGAGATCTATTTCACAAATGTTGAAGACGAAAGGATATACATAAAGACATCTT 909

Qy 101 GlnGluLysLysAspGlySerIleLysProGluGlnLysGluAspLysSer 117
Db 910 GAAGAAAAAGAAAGATGTTCAATTAACCAAGAAACAAATGATTAATCT 960

RESULT 4
US-09-900-963-41
Sequence 41, Application US/09900963
Publication No. US20030064075A1
GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUTHE, Pierre
TITLE OF INVENTION: HEPATIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 41:

US-09-900-963-41

Alignment Scores:
Pred. No.: 7.61e-66 Length: 1482
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-900-963-41 (1-1482)

Qy 1 LeuGInGluGInGInaGAspLeuGluGInaGlyAspIleThrlLysLysAsnLeuGlu 20
|||||

Db 610 TTGCAAGACGACAAAGATTAGAACAAAGAAAGCTGATACGAAAAAATTAGAA 669
Qy 21 ArglyblysglunhiglyaspdlleleuagluaspdeuTyrglyArgleugluilePro 40
Db 670 AAAAAAAGAAACATGAGATATATTAGACAGAGATTATATGCTGTTAGAAAAACCA 729
Qy 41 AlailegluleuProsergluansgluArglyTyrrlyleProhiglnserleu 60
Db 730 GCTATGAACTTCATCAGAAAAATGACGTGATATATATACCATCAATCTCTTTA 789
Qy 61 ProglinaapanaArglyAasnerArgaserysgluileSerlleleglulysThr 80
Db 790 CCTCAGGACAAACAGAGGAGATAGTAGAGATCCAGAAATATCTATATAGAAAAACA 849
Qy 81 AanaArglyserlleThrThrAsnValGluGlyArgArgAspLleHislysglyHisleu 100
Db 850 AATAGAGATCTATTACAACAATGTTGAAGACGAAAGGATATATATTAAGACATCTT 909
Qy 101 GluGluLyblyAspGlySerlleYsProGluGluGluAspLysSer 117
Db 910 GAAGAAAAAGAAAGATGCTTCATTAACCAAGAAACAAAAAGAAATTAATCT 960

RESULT 5
US-09-837-344-38
; Sequence 38, Application US/09837344
; Patent No. US20020041882A1
; GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/837,344
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-837-344-38

Alignment Scores:
Pred. No.: 7,68e-66 Length: 1493
Score: 603.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-837-344-38 (1-1493)

Qy 1 LeuglnglunGlnArgAspLeugluGlnArglyblyAaepThrLyblyAsnleuglu 20
Db 607 TTGCAAGACGACAAAGATTAGAACAAAGAAAGCTGATACGAAAAAATTAGAA 666
Qy 21 ArglyblysglunhiglyaspdlleleuagluaspdeuTyrglyArgleugluilePro 40
Db 667 AAAAAAAGAAACATGAGATATATTAGACAGAGATTATATGCTGTTAGAAAAACCA 726
Qy 41 AlailegluleuProsergluansgluArglyTyrrlyleProhiglnserleu 60
Db 727 GCTATGAACTTCATCAGAAAAATGACGTGATATATATACCATCAATCTCTTTA 786
Qy 61 ProglinaapanaArglyAasnerArgaserysgluileSerlleleglulysThr 80
Db 787 CCTCAGGACAAACAGAGGAGATAGTAGAGATCCAGAAATATCTATATAGAAAAACA 846
Qy 81 AanaArglyserlleThrThrAsnValGluGlyArgArgAspLleHislysglyHisleu 100
Db 847 AATAGAGATCTATTACAACAATGTTGAAGACGAAAGGATATATATTAAGACATCTT 906
Qy 101 GluGluLyblyAspGlySerlleYsProGluGluGluAspLysSer 117
Db 907 GAAGAAAAAGAAAGATGCTTCATTAACCAAGAAACAAAAAGAAATTAATCT 957

RESULT 6
US-09-900-963-38
; Sequence 38, Application US/09900963
; Publication No. US20030064075A1
; GENERAL INFORMATION:
APPLICANT: GUERIN-MARCHAND, Claudine
DRUILHE, Pierre
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE
HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE
OF STIMULATING THE T LYMPHOCYTES
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-900-963-38

Alignment Scores:

Pred. No.:	7,686-66	Length:	1493
Score:	603.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-837-344-20 (1-117) x US-09-900-963-38 (1-1493)

QY 1 LeuGlGluGlnGlnAArgApLeuGluGlnArgLyAlAAspThrLySlySAsnLeuGlu 20
DB 607 TTCGAAGAGCAGCAAGAGATTGAGACAAAGAGAGCTGATACGAAAAATTTAGAA 666
QY 21 ArgLySlySgLuHsGlyAspLleLeuAlaGluAspLeuTyGlyArgLeuGluLlePro 40
DB 667 AGAAAAAGGACACTGAGATATATTAGCAGAGATTATATGCTCTTTAGAAATACCA 726
QY 41 AlaIleGluLeuProSerGluAenGluArgLyTyTrTyleProHsGlnSerLeu 60
DB 727 GCTATAGAACTTCATCAGAAAAAGAACGCGATATTATATCCACATCATCTTCTTTA 786
QY 61 ProGlnAspAsnArgGlyAenSerArgAspSerLySgLuLleSerIleIleGluLyThr 80
DB 787 CCTCAGACAAACAGAGGAATAGTAGAGATTCCAGAGAAATATCTATATAGAAAAACA 846
QY 81 AsnArgLysSerIleThrTrAsnValGluGlyArgAspLleHsLySgLyHsLeu 100
DB 847 AATAGAGATCTTATTACAAACAATGTTGAAGAGGATATATACATAAGACATCTT 906
QY 101 GluGluLySlySAspGlySerIleLyPProGluGlnLySgLuAspLySer 117
DB 907 GAAGAAAAAGAAAGATGTTCAATATAAACCAAGAACAAAAAGATTAATCT 957

RESULT 7

US-09-837-344-45
Sequence 45, Application US/09837344
Patent No. US20020041882A1

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

DRUILHE, Pierre

TITLE OF INVENTION: PEPITIDE SEQUENCES SPECIFIC FOR THE

HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathie

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22133-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/837,344

FILING DATE: 19-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/462,625
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 91 01286
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:

LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-837-344-45

Alignment Scores:

Pred. No.:	1,826-65	Length:	1482
Score:	600.00	Matches:	116
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.15%	Mismatches:	0
Query Match:	99.50%	Indels:	0
DB:	3	Gaps:	0

US-09-837-344-20 (1-117) x US-09-837-344-45 (1-1482)

QY 1 LeuGlGluGlnGlnAArgApLeuGluGlnArgLyAlAAspThrLySlySAsnLeuGlu 20
DB 610 TTCGAAGAGCAGCAAGAGATTGAGACAAAGAGAGCTGATACGAAAAATTTAGAA 669

QY 21 ArgLySlySgLuHsGlyAspLleLeuAlaGluAspLeuTyGlyArgLeuGluLlePro 40
DB 670 AGAAAAAGGACACTGAGATATATTAGCAGAGATTATATGCTCTTTAGAAATACCA 729

QY 41 AlaIleGluLeuProSerGluAenGluArgLyTyTrTyleProHsGlnSerLeu 60
DB 730 GCTATAGAACTTCATCAGAAAAAGAACGCGATATTATATCCACATCATCTTCTTTA 789

QY 61 ProGlnAspAsnArgGlyAenSerArgAspSerLySgLuLleSerIleIleGluLyThr 80
DB 790 CCTCAGACAAACAGAGGAATAGTAGAGATTCCAGAGAAATGCTATATATAGAAAAACA 849

QY 81 AsnArgLysSerIleThrTrAsnValGluGlyArgAspLleHsLySgLyHsLeu 100
DB 850 AATAGAGATCTTATTACAAACAATGTTGAAGAGGAGATATACATAAGACATCTT 909

QY 101 GluGluLySlySAspGlySerIleLyPProGluGlnLySgLuAspLySer 117
DB 910 GAAGAAAAAGAAAGATGTTCAATATAAACCAAGAACAAAAAGATTAATCT 960

RESULT 8

US-09-900-963-45
Sequence 45, Application US/09900963
Publication No. US20030064075A1

GENERAL INFORMATION:

APPLICANT: GUERIN-MARCHAND, Claudine

DRUILHE, Pierre

TITLE OF INVENTION: PEPITIDE SEQUENCES SPECIFIC FOR THE

HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE

OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46

```

CORRESPONDENCE ADDRESS:
ADDRESSSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/900,963
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/098,327
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO 92/13884
PUBLICATION DATE: 20-AUG-1992
SEQUENCE DESCRIPTION: SEQ ID NO: 45:

US-09-900-963-45

Alignment Scores:
Pred. No.: 1,826-65 Length: 1482
Score: 600.00 Matches: 116
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.15% Mismatches: 0
Query Match: 99.50% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-900-963-45 (1-1482)
QY 1 LeuGlnGlnGlnGlnArgAspLeuGlnArgGlyAlaAspThrIysAsnLeuGlu 20
DB 610 TTGCAAGACGACGAAAGAGATTAGAAACAAAGAGAGCTGATTCGAAAAAATTAGAA 669
QY 21 ArgGlyAspGluHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIlePro 40
DB 670 AGAAAAAGGAAACATGAGATATATTAGCAGAGATTATATGCTGTTAGAAATACCA 729
QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerSerLeu 60
DB 730 GCTATGAACCTTCATCAGAAATGAAACGATATTATATACACATCAATCTTCTTA 789
QY 61 ProGluAspAsnArgGlyAsnSerArgAspSerGlyLeuIleSerIleIleGluIysThr 80
DB 790 CCTCAGACACACAGAGGAAATAGTAGATTCCAGGAAATGCTCTATATAGAAAAACA 849
QY 81 AsnArgGluSerIleThrIleAsnValGluGlyArgArgAspIleHisIysGlyHisLeu 100
DB 850 AATAGAGAAATCTATTCAACAATAGTTGAAAGAGCAAGAGGATATACATTAAGACATCTT 509
QY 101 GluGluGlyAspGlySerIleIleProGluGlnGluAspIysSer 117
```

```

DB 910 GAAGAAAAAGAAAGATGTTCAATAAACCAAGACAAAAAGAGATTAATCT 960

RESULT 9
US-09-894-018-122
Sequence 122, Application US/09894018
Patent No. US20020119127A1
GENERAL INFORMATION:
APPLICANT: EPIMUNE, Inc.
APPLICANT: Sette, Alessandro
APPLICANT: Chestnut, Robert
APPLICANT: Livingston, Brian
APPLICANT: Baker, Denise
APPLICANT: Newman, Mark
APPLICANT: Brown, David
TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
FILE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
FILE REFERENCE: 39963-2003..00
CURRENT APPLICATION NUMBER: US/09/894,018
CURRENT FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: PCT/US00/35568
PRIOR FILING DATE: 2000-12-28
PRIOR APPLICATION NUMBER: US 60/173,390
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: US 60/284,221
PRIOR FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 368
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 122
LENGTH: 1385
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Encoding sequence for PCT/HTL(N)
US-09-894-018-122

Alignment Scores:
Pred. No.: 0,161 Length: 1385
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: 3 Gaps: 0

US-09-837-344-20 (1-117) x US-09-894-018-122 (1-1385)
QY 40 ProAlaIleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerSer 59
DB 361 CCTGACACAGGCTGCTCCGTCGAAACGAAAGAGGATACATACCTCACAGAGCAGC 420
QY 60 Leu 60
DB 421 CTC 423

RESULT 10
US-10-474-960A-122
Sequence 122, Application US/10474960A
Publication No. US20040248113A1
GENERAL INFORMATION:
APPLICANT: Sette, Alessandro
APPLICANT: Chestnut, Robert
APPLICANT: Livingston, Brian
APPLICANT: Baker, Denise
APPLICANT: Newman, Mark
APPLICANT: Brown, David
TITLE OF INVENTION: Methods and System for Optimizing Multi-epitope Nucleic
FILE OF INVENTION: Acid Constructs and Peptides Encoded Thereby
FILE REFERENCE: 2060.0320004
CURRENT APPLICATION NUMBER: US/10/474,960A
CURRENT FILING DATE: 2003-10-16
PRIOR APPLICATION NUMBER: PCT/US02/09877
PRIOR FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 09/894,018
```

```

; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encoding sequence for PECTL/HTL(N)
US-10-474-960A-122

Alignment Scores:
Pred. No.: 0.161 Length: 1385
Score: 94.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 15.59% Indels: 0
DB: Gaps: 0

US-09-837-344-20 (1-117) x US-10-474-960A-122 (1-1385)

QY 40 ProbableLeuProSerGluAsnGluArgGlyTyrIleProHisGlnSerSer 59
DB 361 CTCGACCAAGCTCTGCGTCCGAAACGAAAGAGATACATACCTCCACGAGCAGC 420

QY 60 Leu 60
DB 421 CTC 423

RESULT 11
US-10-450-763-7484/c
; Sequence 7484, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 7484
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (482)..(183)
; OTHER INFORMATION: 51% homologous to Homo sapiens p40, accession number
; OTHER INFORMATION: U93569, Smith-Waterman Score=226.
US-10-450-763-7484

Alignment Scores:
Pred. No.: 0.28 Length: 969
Score: 90.50 Matches: 30
Percent Similarity: 47.66% Conservative: 21
Best Local Similarity: 28.04% Mismatches: 35
Query Match: 15.01% Indels: 21
DB: Gaps: 6

US-09-837-344-20 (1-117) x US-10-450-763-7484 (1-969)

QY 1 LeuGlnGluGlnGlnArgAspLeuGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
DB 600 CTCACGACGCTCCACAGACCTGACGCTGACCTGACTGTTAGAGAAAACTAAC 541

QY 600 CTCACGACGCTCCACAGACCTGACGCTGACCTGACTGTTAGAGAAAACTAAC 541
```

```

QY 21 ArgLysLysGlnHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGlnIlePro 40
DB 540 AAACAGAAAGAGACATCCACAG-----CAAAACCC 511

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyr-----IleProHisGln 57
DB 510 ATCTGTACGTCAACATCATCAAAAGACCAAGATATTATCCAGAGAACTTCCCAACT 451

QY 58 SerSerLeuProGln---AspAsnArgGlyAsnSerArgAspSerLysGlnIleSerIle 76
DB 450 AGCAAGACAGGCCAGCATTCAAATTCCAGGAATATCACAGAAATCCCAAGATATA---CTC 394

QY 77 IleGluLysThrAsnArgLysSerIleThrThrAsnValGluGlyArgArgAspIleHis 96
DB 393 CTCGAGAAAGAGCAAC-----TCCAAAGACATATATTGTCAG-----ATTCCAC 352

QY 97 LysGlyHisLeuGlnGluLys 103
DB 351 CAACTTGAAATGAAGAAAAA 331

RESULT 12
US-10-450-763-11162/c
; Sequence 11162, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 11162
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (200)..(430)
; OTHER INFORMATION: 100% homologous to Homo sapiens proteasome subunit
; OTHER INFORMATION: XAPC7, accession number AF022815, Smith-Waterman Score=404.
US-10-450-763-11162

Alignment Scores:
Pred. No.: 0.28 Length: 969
Score: 90.50 Matches: 30
Percent Similarity: 47.66% Conservative: 21
Best Local Similarity: 28.04% Mismatches: 35
Query Match: 15.01% Indels: 21
DB: Gaps: 6

US-09-837-344-20 (1-117) x US-10-450-763-11162 (1-969)

QY 1 LeuGlnGluGlnGlnArgAspLeuGlnArgLysAlaAspThrLysLysAsnLeuGlu 20
DB 600 CTCACGACGCTCCACAGACCTGACGCTGACCTGACTGTTAGAGAAAACTAAC 541

QY 21 ArgLysLysGlnHisGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGlnIlePro 40
DB 540 AAACAGAAAGAGACATCCACAG-----CAAAACCC 511

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyr-----IleProHisGln 57
DB 510 ATCTGTACGTCAACATCATCAAAAGACCAAGATATTATCCAGAGAACTTCCCAACT 451

QY 58 SerSerLeuProGln---AspAsnArgGlyAsnSerArgAspSerLysGlnIleSerIle 76
```

```
Db 450 ACCAAGACAGCGCGATTCATAATTCAGAAATTCACAGAAATGCCAAGATTA---CTC 394
Qy 77 lIeGluYtHrAsnArgLseRlIeThrAsnValGluGlyArgArgAspIleHis 96
Db 393 CTCGAGAAAGCAGAAC-----TCCAGACACATTAATTGTTCAG-----ATTGAC 352
Qy 97 LyvGlyHlIeGluGluGly 103
Db 351 CAAGTTGAATGAGAGAAAA 331

RESULT 13
US-10-450-763-28063
; Sequence 28063, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 28063
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (488) ..(787)
; OTHER INFORMATION: 51% homologous to Homo sapiens p40, accession number
; OTHER INFORMATION: U93569, Smith-Waterman Score=226.
US-10-450-763-28063

Alignment Scores:
Pred. No.: 0.28 Length: 969
Score: 90.50 Matches: 30
Percent Similarity: 47.66% Conservative: 21
Best Local Similarity: 28.04% Mismatches: 35
Query Match: 15.01% Indels: 21
DB: 9 Gaps: 6

US-09-837-344-20 (1-117) x US-10-450-763-28063 (1-969)
Qy 1 lEuglInGluInGlnArgAspLeuGluGlnArgLysAlaAspThrLysValAsnLeuGlu 20
Db 370 CTCGACGAGAGCTCCACAGACCTGACGTGAGGATCCTGACTGTGAGAGAAATTAAC 429
Qy 21 ArgLysLysGluHlIeGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGluIlePro 40
Db 430 AAACAGAAAGACATCCACAG-----CAAAACCC 459
Qy 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyr-----IleProHlIeGln 57
Db 460 ATCTGTAACCTCACCATCATCAAGACCAAGATATTATCCAGAGAACTCCCAACCT 519
Qy 58 SerSerLeuProGln--AspAsnArgGlyAsnSerArgAspSerLysGluIleSerIle 76
Db 520 AGCAAGACAGCGCAGCATTCATAATTCCAGAAATACACAGAAATGCCAAGATATA---CTC 576
Qy 77 lIeGluYtHrAsnArgLseRlIeThrAsnValGluGluGlyArgArgAspIleHis 96
Db 577 CTCGAGAAAGCAGAAC-----TCCAGACACATTAATTGTTCAG-----ATTGAC 618
Qy 97 LyvGlyHlIeGluGluGly 103
Db 619 CAAGTTGAATGAGAGAAAA 639
```

```
RESULT 14
US-10-172-118-1520
; Sequence 1520, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1520
; LENGTH: 4080
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_014700
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1520

Alignment Scores:
Pred. No.: 2.1 Length: 4080
Score: 90.00 Matches: 31
Percent Similarity: 41.67% Conservative: 24
Best Local Similarity: 23.48% Mismatches: 53
Query Match: 14.93% Indels: 24
DB: 6 Gaps: 4

US-09-837-344-20 (1-117) x US-10-172-118-1520 (1-4080)
Qy 2 GInGluInGlnArgAspLeuGluGlnArgLysAlaAspThrLysValAsnLeuGluArg 21
Db 2167 CAGGAGAACAGAGAGAGATGGGGAGCAGAGCTGAGTCAAGAGCACCAGTTCCAGAGG 2226
Qy 22 LysLysGluHlIeGlyAspIleLeuAlaGluAspLeuTyrGlyArgLeuGlu---IlePro 40
Db 2227 GACAAAGAGGCCACCCAGAGAGCTGATCGAGGACCTCCAAAGCAGCTGAGCACTGCAG 2286
Qy 41 AlaIleGluLeuProSerGluAsnGluArgGly----- 51
Db 2287 CTCCTCAAGCTGAGAGCCGAGCGCGGGCCCGCAGAGCAGCATGGGCTGCAGAGG 2346
Qy 52 TyrTyrIleProHlIeGlnSerSer-----LeuProGlnAsp 63
Db 2347 TACCAAGCGCGCGCGGAGAGAGAGCTGAGCAGAGAGTCCGAGGCTGAAGCAGAGAC 2406
Qy 64 AsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluYtHrAsnArgGlu 83
Db 2407 AACCCCAACCTTAAGAGCAGAAAGAGAGCTG-----AACGGGCGAG 2448
Qy 84 SerIleThrAsnValGluGluArgArgAspIleHisLysGlyHlIeGluGluGly 103
Db 2449 ATCATTTACCTCAGCATTCAGAGGCCCAAGACCTTTCTCCAGAGCCTTCTGTAGTCC 2508
Qy 104 LyvAspGlySerIleLysProGluGluGluArg 115
Db 2509 CTCGCTGCAGAGATCAGCTCCGCTCCGAGATGAG 2544

RESULT 15
US-10-342-887-1520
; Sequence 1520, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernard, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIORITY FILING DATE: 2003-01-15
; PRIORITY FILING DATE: 2001-06-18
; PRIORITY FILING DATE: 2001-06-18
; PRIORITY FILING DATE: 2002-05-14
; PRIORITY FILING DATE: 2002-05-14
; PRIORITY FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1520
; LENGTH: 4080
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1520

Alignment Scores:
Pred. No.: 2.1          Length: 4080
Score: 90.00          Matches: 31
Percent Similarity: 41.67%      Conservative: 24
Best Local Similarity: 23.48%    Mismatches: 53
Query Match: 14.93%            Indels: 24
DB: 7                        Gaps: 4

US-09-837-344-20 (1-117) x US-10-342-887-1520 (1-4080)

QY      2  GlnGluGlnGlnArgAspLeuGluGlnArgLysAlaSerThrLysLysAsnLeuGluArg 21
      |||||  ::::|  ::::|  |||  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
Db      2167  CAGAGAGACAGAGAGAGAGATGGGAGACGCTGAGTCCAGAGGACACCACTTCCAGAGG 2226

QY      22  LysLysGluHISGlyAspLysLeuAlaGluAspLeuTyrgLysArgLeuGlu---IlePro 40
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ::
Db      2227  GACAAAGAGGCCACCCAGAGCTGATCGAGGACCTCCGAAGCAGCTGGAGCAGCTGCAG 2286

QY      41  AlaIleGluLeuProSerGluAsnGluArgGly----- 51
      ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::
Db      2287  CTCCTCAAGCTGAGAGCCGAGCAGCGCGGCGCGCAGACAGACATGCGCTGCAGAGAG 2346

QY      52  TyrTyrIleProHISGlnSerSer-----LeuProGluAsp 63
      |||||  ::::|  ::::|  |||||  |||||  |||||  |||||  |||||  |||||  ::
Db      2347  TACCACAGCCCGCCCGGAGAGCAGCTGAGCAGAGAGTCCGAGGCTGAAGCAGGAGAC 2406

QY      64  AsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThrAsnArgGlu 83
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ::
Db      2407  AACCGCAACTGAAGAGACAGAAACGAGAGCTG-----AACGGCAG 2448

QY      84  SerIleThrAsnValGluGlyArgAspIleHisLysGlyHisLeuGluGluLys 103
      |||||  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::
Db      2449  ATCAATTACCTCAGAGATCCAGAGGCGCAAGAGCTCTTCTCCACAGCTTCTTGAGTCC 2508

QY      104  LysAspGlySerIleLysProGluGlnLysGluAsp 115
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ::
Db      2509  CTGGCTGAGAGATCAGCTCCGCTCCGAGATGAG 2544
```

Search completed: December 7, 2005, 16:14:58
Job time : 414.905 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 7, 2005, 12:39:12 ; Search time 80.5219 Seconds
(without alignments)
543.218 Million cell updates/sec

Title: US-09-837-344-20

Perfect score: 603 1 LOGOQDRLEQRKADTKKNLR.....GHLEKKDSIKRQKEDKS 117

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgn2 1/USPTO.spool h/US09837344/runat 05122005 133535 865/app.query.fasta_1.718
-DB=Published Applications NA New -OPMT=fastap -SUFFIX=rnbn -MINMATCH=0.1
-LOOPC=0 -LOGEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09837344 @CIGN 1 1 231 @runat 05122005 133535 865
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	86	14.3	1344	US-10-793-626-2899 Sequence 2899, Ap
2	86	14.3	1353	US-10-793-626-3225 Sequence 3225, Ap
3	86	14.3	3001	US-10-793-626-3813 Sequence 3813, Ap
4	86	14.3	3974	US-10-793-626-4245 Sequence 4245, Ap
5	82.5	13.7	193084	US-11-121-086-82 Sequence 82, Appl
6	79.5	13.2	8424	US-10-821-234-47 Sequence 47, Appl
7	79	13.1	423	US-10-793-626-2267 Sequence 2267, Ap
8	79	13.1	423	US-10-793-626-2437 Sequence 2437, Ap

9	79	13.1	2525	US-10-770-726-5	Sequence 5, Appl1
10	79	13.1	3019	US-10-793-626-4232	Sequence 4232, Ap
11	79	13.1	3726	US-10-793-626-3917	Sequence 3917, Ap
12	79	13.1	3986	US-10-793-626-4014	Sequence 4014, Ap
13	78.5	13.0	2040	US-10-750-185-41854	Sequence 41854, A
14	77.5	12.9	120697	US-11-121-086-48	Sequence 48, Appl
15	77	12.8	173602	US-11-121-086-25	Sequence 25, Appl
16	75	12.4	1458	US-10-508-263-21	Sequence 21, Appl
17	75	12.4	8948	US-10-623-155-119	Sequence 119, Ap
18	75	12.4	150468	US-11-112-908-56	Sequence 56, Appl
19	75	12.4	193789	US-11-112-908-55	Sequence 55, Appl
20	74	12.3	38862	US-11-121-086-76	Sequence 76, Appl
21	72.5	12.0	1476	US-10-793-626-2807	Sequence 2807, Ap
22	72.5	12.0	3857	US-10-793-626-3480	Sequence 3480, Ap
23	71.5	11.9	636	US-10-467-657-5485	Sequence 5485, Ap
24	71.5	11.9	636	US-10-467-657-7403	Sequence 7403, Ap
25	71.5	11.9	648	US-10-793-626-595	Sequence 595, App
26	71.5	11.9	648	US-10-793-626-2257	Sequence 2257, Ap
27	71	11.8	984	US-10-467-657-6641	Sequence 6641, Ap
28	71	11.8	984	US-10-467-657-7709	Sequence 7709, Ap
29	71	11.8	1605	US-10-750-185-60116	Sequence 60116, A
30	71	11.8	98862	US-11-121-086-76	Sequence 76, Appl
31	70.5	11.7	451	US-11-108-172-1038	Sequence 1038, Ap
32	70.5	11.7	1200	US-10-131-826A-273	Sequence 273, App
33	70.5	11.7	2158	US-10-909-125-805	Sequence 805, App
34	70.5	11.7	149419	US-11-112-908-49	Sequence 49, Appl
35	70.5	11.7	161726	US-11-112-908-48	Sequence 48, Appl
36	70.5	11.7	161726	US-11-112-908-52	Sequence 52, Appl
37	70.5	11.7	166111	US-11-112-908-47	Sequence 47, Appl
38	70	11.6	762	US-10-467-657-3449	Sequence 3449, Ap
39	69	11.4	2845	US-10-909-125-1740	Sequence 1740, Ap
40	69	11.4	5233	US-11-060-914-1	Sequence 1, Appl1
41	69	11.4	162537	US-11-121-086-59	Sequence 59, Appl
42	68.5	11.4	207835	US-11-121-086-39	Sequence 39, Appl
43	68.5	11.4	207835	US-11-121-086-40	Sequence 40, Appl
44	68	11.3	821	US-10-750-185-46312	Sequence 46312, A
45	68	11.3	2583	US-10-750-185-57927	Sequence 37927, A

ALIGNMENTS

RESULT 1
US-10-793-626-2899
; Sequence 2899, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P33480US
; CURRENT APPLICATION NUMBER: US/10/793, 626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164, 258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2899
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2899

Alignment Scores:

Pred. No.: 1.21 Length: 1344
Score: 86.00 Matches: 29
Percent Similarity: 43.44% Conservative: 24
Best Local Similarity: 23.77% Mismatches: 43
Query Match: 14.26% Indels: 26
DB: Gaps: 4
US-09-837-344-20 (1-117) x US-10-793-626-2899 (1-1344)

```
QY      1 LeuGlnGlnGlnGlnArgAsePleuGlnGlnArgLyAlaAsePThrLyLeuAsePLeuGlu 20
      43 ATTCTAAAGCGCAAGAGAGCTGCAAAAAGCGCACTTCTGAAAGCTCTGTAATTTAA 102
QY      21 -----ArgLyLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 35
      103 GCTGATATTGATGCTCAAAAAGAGATGCAAGAACTC-----GAACAGCTTTCAAAA 156
QY      36 ArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 55
      157 GAAATTGAAGCGCTCAACA-----CCT 177
QY      56 HisGlnSerSerLeuProGlnAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 75
      178 AAACAAGATGAACCACTTAAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 237
QY      76 IleIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 95
      238 TCTGAGAAAGAAATCAGAGAAACCTTCTGATGATGAACCAAGAAAGAACTTCAAGT--- 294
QY      96 HisLyGlnHisLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 115
      295 -----GAAGAAAACCTGATGATGACCAAAACCAAGATGACAAACCTGAA 339
QY      116 LysSer 117
      340 GAAACA 345
DB
RESULT 2
US-10-793-626-3225
/ Sequence 3225, Application US/10793626
/ Publication No. US20050255478A1
/ GENERAL INFORMATION:
/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PU3480US
/ CURRENT APPLICATION NUMBER: US/10/793,626
/ CURRENT FILING DATE: 2004-03-04
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3225
/ LENGTH: 1353
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
/ OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3225

Alignment Scores:
Pred. No.: 1 22 Length: 1353
Score: 86.00 Matches: 29
Percent Similarity: 43.44% Conservative: 24
Best Local Similarity: 23.77% Mismatches: 43
Query Match: 14.26% Indels: 26
DB: 6 Gaps: 4

US-09-837-344-20 (1-117) x US-10-793-626-3225 (1-1353)
QY      1 LeuGlnGlnGlnGlnArgAsePleuGlnGlnArgLyAlaAsePThrLyLeuAsePLeuGlu 20
      43 ATTCTAAAGCGCAAGAGAGCTGCAAAAAGCGCACTTCTGAAAGCTCTGTAATTTAA 102
QY      21 -----ArgLyLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 35
      103 GCTGATATTGATGCTCAAAAAGAGATGCAAGAACTC-----GAACAGCTTTCAAAA 156
QY      36 ArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 55
      157 GAAATTGAAGCGCTCAACA-----CCT 177
DB
157 GAAATTGAAGCGCTCAACA-----CCT 177
```

```
QY      56 HisGlnSerSerLeuProGlnAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 75
      178 AAACAAGATGAACCACTTAAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 237
QY      76 IleIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 95
      238 TCTGAGAAAGAAATCAGAGAAACCTTCTGATGATGAACCAAGAAAGAACTTCAAGT--- 294
QY      96 HisLyGlnHisLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 115
      295 -----GAAGAAAACCTGATGATGACCAAAACCAAGATGACAAACCTGAA 339
QY      116 LysSer 117
      340 GAAACA 345
DB
RESULT 3
US-10-793-626-3813/c
/ Sequence 3813, Application US/10793626
/ Publication No. US20050255478A1
/ GENERAL INFORMATION:
/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PU3480US
/ CURRENT APPLICATION NUMBER: US/10/793,626
/ CURRENT FILING DATE: 2004-03-04
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3813
/ LENGTH: 3001
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
/ OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3813

Alignment Scores:
Pred. No.: 3 44 Length: 3001
Score: 86.00 Matches: 29
Percent Similarity: 43.44% Conservative: 24
Best Local Similarity: 23.77% Mismatches: 43
Query Match: 14.26% Indels: 26
DB: 6 Gaps: 4

US-09-837-344-20 (1-117) x US-10-793-626-3813 (1-3001)
QY      1 LeuGlnGlnGlnGlnArgAsePleuGlnGlnArgLyAlaAsePThrLyLeuAsePLeuGlu 20
      2456 ATTCTAAAGCGCAAGAGAGCTGCAAAAAGCGCACTTCTGAAAGCTCTGTAATTTAA 2397
QY      21 -----ArgLyLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 35
      2396 GCTGATATTGATGCTCAAAAAGAGATGCAAGAACTC-----GAACAGCTTTCAAAA 2343
QY      36 ArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 55
      2342 GAAATTGAAGCGCTCAACA-----CCT 2322
QY      56 HisGlnSerSerLeuProGlnAsePLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 75
      2321 AAACAAGATGAACCACTTAAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2262
QY      76 IleIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 95
      2261 TCTGAGAAAGAAATCAGAGAAACCTTCTGATGATGAACCAAGAAAGAACTTCAAGT--- 2205
QY      96 HisLyGlnHisLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 115
      2204 -----GAAGAAAACCTGATGATGACCAAAACCAAGATGACAAACCTGAA 2160
DB
```

QY 116 LySer 117
DB 2159 GAACA 2154

RESULT 4
US-10-793-626-4245
Sequence 4245, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P134805
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4245
LENGTH: 3974
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4245

Alignment Scores:
Pred. No.: 4.97 Length: 3974
Score: 86.00 Matches: 29
Percent Similarity: 43.44% Conservative: 24
Best Local Similarity: 23.77% Mismatches: 43
Query Match: 14.26% Indels: 26
DB: 6 Gaps: 4

US-09-837-344-20 (1-117) x US-10-793-626-4245 (1-3974)

QY 1 LeuGlnGluGlnGlnArgSerLeuGlnGlnArgLysAlaSerThrLysLysAsnLeuGlu 20
DB 460 ATTCTAAGCGCAAGAAAGACGCGAAGGCGACCTCGAACTGCTGTAATTAAAA 519
QY 21 -----ArgLysLysGlnGlnGlnArgLysAlaSerThrLysLysAsnLeuGlu 35
DB 520 GCTGATATGTGATGCTCAAAAGAAAGAGATGAGAAAGAACCTC-----GAACAGCTTTCAAAA 573
QY 36 ArgLeuGluLeuProAlaLeuGluLeuProSerGluAsnGluArgLysLysLysPro 55
DB 574 GAATTGAAAGCGTCAGCA-----CCT 594
QY 56 HisGlnSerSerLeuProGlnAspAsnArgLysAsnSerArgSerLysGluLysSer 75
DB 595 AAACAGATGAAACCACTAAAGATGAAAGCTGACAGAGTTGAAGTAACAAAGATGTAAT 654
QY 76 IleLeuGluLysThrAsnArgLysSerLysThrAsnValGluGlyArgArgAspLys 95
DB 655 TCTGAGAGAAATGACGAAACAACTTCTGATGATGAAACCAAGAGAACTTCAGAT--- 711
QY 96 HisLysGluHisLysGluGlnLysLysAspLysSerLysLysProGlnGlnLysGluAsp 115
DB 712 -----GAAGAAAGAACTGATGATGACCAAAACAGATGACAAACCTGAA 756
QY 116 LySer 117
DB 757 GAACA 762

RESULT 5
US-11-121-086-82
Sequence 82, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KRISTEN V.

TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 82
LENGTH: 193084
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-82

Alignment Scores:
Pred. No.: 2.12e+03 Length: 193084
Score: 82.50 Matches: 35
Percent Similarity: 43.18% Conservative: 22
Best Local Similarity: 26.52% Mismatches: 44
Query Match: 13.68% Indels: 31
DB: 7 Gaps: 7

US-09-837-344-20 (1-117) x US-11-121-086-82 (1-193084)

QY 1 LeuGlnGlnGlnGlnArgSerLeuGlnGlnArgLysAlaSerThrLysLysAsnLeuGlu 20
DB 137155 ATATCAGAGATTGGAAGTGAATCTTAATGAAATAGAGAGCAAAATTTAGAAAAA 137214
QY 21 ArgLysLysGlnGlnGlnArgSerLysAlaSerThrLysLysAsnLeuGlu 38
DB 137215 AGAATGAAAGAAACAACAAGCTCCAGAAATATGAGCTATGAAAGCAAAAT 137274
QY 39 IleProAlaLeuGluLeuProSer-----GluAsnGlu 49
DB 137275 TTACGTTTGAATGGAATATCCTGAAAGTGAAGAGATGAAACCAAGTTGAAACTCT 137334
QY 50 ArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 62
DB 137335 TCAGGATATTTAT-----CCAGGAACTTCCACAACTTGAAGAGACAGCCAAAT 137385
QY 63 AspAsnArgLysAsnSerArgSerLysGluLysSerLysLysLysLysLysLysLysLys 82
DB 137386 TCAATTAAAGGAATATACAGAAACCAACAAGCTA---CTCCGAGGAAGAGCAACCC 137442
QY 83 GluSerLysThrAsnValGluGlyArgArgAspLysHisLysGlyHisLysGluGlu 102
DB 137443 AAGACA-----CATATCATGATGATTCACCAAGGTCAAAATGAAGA 137484
QY 103 LysLysAspGlySerLysLysProGlnGlnLysGlu 114
DB 137485 AAA-----AATGTTAAGGCGACCGCAGAGAA 137511

RESULT 6
US-10-821-234-47
Sequence 47, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmani, Susan
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PL_SEQ_genes Version 1.0
SEQ ID NO 47
LENGTH: 8424
TYPE: DNA
ORGANISM: Homo sapiens

US-10-821-234-47

Alignment Scores:

Alignment No.:	82.8	Length:	8424
Score:	79.50	Matches:	29
Percent Similarity:	42.20%	Conservative:	17
Best Local Similarity:	26.61%	Mismatches:	38
Query Match:	13.18%	Indels:	25
DB:	6	Gaps:	5

US-09-837-344-20 (1-117) x US-10-821-234-47 (1-8424)

QY 17 LysAsnLeuGluArgLysLeu-----HisGlyAspIleLeuAlaGlu 31

DB 5267 AAGCAGTGAACGCTTTCAAGACCTTAGACCTGCATTGAAGACCTTACTTCAAGACT 5326

QY 32 AspLeuTyrGlyArgLeuGluIleProAlaIleGluLeuProSerGluAsnGluArgGly 51

DB 5327 GATGTT--GGAAGT-----GACAGATCT 5347

QY 52 TyrTyrIleProHisIleGlnSerLeuProGln-----AspAsnArgGlyAsnSer 68

DB 5348 GCTGCCAGCCCGGTTGTAAGTATGACGAAAGAGCTTCAGAGCTTCTTCAGAGCAA 5407

QY 69 ArgAspSerLysGluIleSerIleIleGluLysThrAsnArgGluSerIleThrAsn 88

DB 5408 AAGAGATGATTATGAAATTTTGTAAAGTTAAGACACTCAGCAAAAAGCAAGAAAT 5467

QY 89 ValGluGlyArgArgAspIleHisLysGlyHisLeuGluGluLysAspGlySerIle 108

DB 5468 -----AAGAACCGTGATTAAGGGGAGAAAGAGAAAGAGACTTCTCATTA 5515

QY 109 LysProGluGlnLysGluAspLysSer 117

DB 5516 AGATCTCGAAGTAAAGCGTTCCAATCT 5542

RESULT 7

US-10-793-626-2267

Sequence 2267, Application US/10793626

Publication No. US20050255478A1

GENERAL INFORMATION:

APPLICANT: KIMBERLY, WILLIAM JOHN

TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

FILE REFERENCE: PU3480US

CURRENT APPLICATION NUMBER: US/10/793,626

CURRENT FILING DATE: 2004-03-04

PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 4472

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2267

LENGTH: 423

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-2267

Alignment Scores:

Pred. No.:	1.92	Length:	423
Score:	79.00	Matches:	28
Percent Similarity:	46.40%	Conservative:	30
Best Local Similarity:	22.40%	Mismatches:	35
Query Match:	13.10%	Indels:	32
DB:	6	Gaps:	6

US-09-837-344-20 (1-117) x US-10-793-626-2267 (1-423)

QY 1 LeuGlnGluGlnArgAspLeuGluGlnArgLysAlaAspThrLysLysAsnLeuGlu 20

DB 100 TTGAAAAAAGAAATTAAGATTATAAAAAACAGAAAGAAAAAGTTACAAAGCAAAAGAT 159

QY 21 ArgLysLysGluHisGlyAspIleLeuAlaGluLysLeuTyrGlyArgLeuGluIlePro 40

DB 160 AAGCTTAAACAAACAGAGATAGCTCCAGAAAGATGTAAT----- 201

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerLeu 60

DB 202 -----GACTTGCCCTTAAGAAC-----ACATCCCGA 228

QY 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80

DB 229 GATTAAGAAAAATTAAGATTAAT--CATATGCAAAAAGAAAGCTTCAGATATCAATGC 285

QY 81 -----AsnArgGluSerIleThrThrAsnValGluGlyArgArgAspIleHis-- 96

DB 286 ACATCTGCTAATCATGATGATCAAACTTAACAAATTAAGCAATCAAGATCAATGCAC 345

QY 97 -----LysGlyHisLeuGluGluLysLysAspGlySerIleLysProGluGln 112

DB 346 AGTCAATCCTCTTAACCAATACATACAGCAG-----AAGCCTTCACAG 387

QY 113 LysGluAspLysSer 117

DB 388 AATGATAGAAAAAT 402

RESULT 8

US-10-793-626-2437

Sequence 2437, Application US/10793626

Publication No. US20050255478A1

GENERAL INFORMATION:

APPLICANT: KIMBERLY, WILLIAM JOHN

TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

FILE REFERENCE: PU480US

CURRENT APPLICATION NUMBER: US/10/793,626

CURRENT FILING DATE: 2004-03-04

PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 4472

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2437

LENGTH: 423

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-2437

Alignment Scores:

Pred. No.:	1.92	Length:	423
Score:	79.00	Matches:	28
Percent Similarity:	46.40%	Conservative:	30
Best Local Similarity:	22.40%	Mismatches:	35
Query Match:	13.10%	Indels:	32
DB:	6	Gaps:	6

US-09-837-344-20 (1-117) x US-10-793-626-2437 (1-423)

QY 1 LeuGlnGluGlnArgAspLeuGluGlnArgLysAlaAspThrLysLysAsnLeuGlu 20

DB 100 TTGAAAAAAGAAATTAAGATTATAAAAAACAGAAAGAAAAAGTTACAAAGCAAAAGAT 159

QY 21 ArgLysLysGluHisGlyAspIleLeuAlaGluLysLeuTyrGlyArgLeuGluIlePro 40

DB 160 AAGCTTAAACAAACAGAGATAGCTTCAGAAAGATGTAAT----- 201

QY 41 AlaIleGluLeuProSerGluAsnGluArgGlyTyrTyrIleProHisGlnSerLeu 60

DB 202 -----GACTTGCCCTTAAGAAC-----ACATCCCGA 228

QY 61 ProGlnAspAsnArgGlyAsnSerArgAspSerLysGluIleSerIleIleGluLysThr 80

DB 229 GATTAAGAAAAATTAAGATTAAT--CATGATGCAAAAAGAAAGCTTCAGATATCAATGC 285

Alignment Scores:	35.9	Length:	398
Pred. No.:		Matches:	28
Score:	79.00		

US-09-837-344-20 (1-117) X US-10-750-185-41854 (1-2040)

QY 6 ArgAspLeuGlnGlnArgLysAlaSerPheLysValSerLeuGluArgLysValGluHis 25

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2005, 04:22:54 ; Search time 1139.77 Seconds
(without alignments)
8665.841 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 1482

Sequence: 1 CAAGAACAAACAAGCATCT.....ATATTATGAACTATTA 1482

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: geneeqn1980s:*
- 2: geneeqn1990s:*
- 3: geneeqn2000s:*
- 4: geneeqn2001s:*
- 5: geneeqn2002s:*
- 6: geneeqn2003s:*
- 7: geneeqn2004s:*
- 8: geneeqn2005s:*
- 9: geneeqn2006s:*
- 10: geneeqn2007s:*
- 11: geneeqn2008s:*
- 12: geneeqn2009s:*
- 13: geneeqn2010s:*
- 14: geneeqn2011s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1482	100.0	1496	2	AAQ28119
2	950	64.1	950	2	AAQ28115
3	853.2	57.6	1320	2	AAQ80916
4	586.4	39.6	1566	2	AAQ05140
5	503.4	34.0	1371	12	ADO21941
6	493.6	33.3	1374	12	ADO21919
7	472.8	31.9	988	2	AAQ28117
8	233.4	15.7	1300	12	ADP85917
9	225.8	15.2	3399	2	AAQ05868
10	183.2	12.4	210	1	AAH80899
11	180.2	12.2	1000	12	ADO62833
12	180.2	12.2	1000	12	ADO62832
13	179.4	12.1	1686	2	AAQ87587
14	176.6	11.9	3579	3	AAH70099
15	170.6	11.5	1998	3	AAH70212
16	151.2	10.2	6292	4	AAH46735
17	149.2	10.1	6767	4	AAH46608
18	146	9.9	3662	4	ABJ05274
19	139.4	9.4	9539	4	AAH45347

C	20	139.4	9.4	9539	6	ABK28180
C	21	137.8	9.3	6668	6	ABJ33697
C	22	136.4	9.2	6644	2	AAH33181
C	23	136.4	9.2	7372	2	AAH33182
C	24	136.4	9.2	7797	2	AAH33180
C	25	136.4	9.2	7996	2	AAH33184
C	26	134.2	9.1	5493	13	ADH89675
C	27	133.8	9.0	778	6	ABJ15588
C	28	132.8	9.0	778	6	ABJ15589
C	29	131.4	8.9	5493	13	ADH89401
C	30	131.4	8.9	7442	4	AAH46686
C	31	129.8	8.8	1039	14	ACI64817
C	32	129.8	8.8	1762	6	ABQ20939
C	33	129.8	8.8	1762	6	ABQ20938
C	34	129.2	8.7	1995	6	ABQ20942
C	35	129.2	8.7	1995	6	ABQ20943
C	36	129	8.7	1200	6	ABQ39210
C	37	129	8.7	7597	6	ABJ33011
C	38	129	8.7	7597	6	ABJ33013
C	39	129	8.7	14006	6	ABJ33958
C	40	127.4	8.6	3683	8	ABJ10199
C	41	127	8.6	3683	8	ABJ10053
C	42	123.4	8.3	3037	6	ABQ54066
C	43	123.4	8.3	3037	6	ABQ54067
C	44	123	8.3	12337	6	ABJ34358
C	45	120.8	8.2	612	14	ACI64894

ALIGNMENTS

RESULT 1	AAQ28119	standard; DNA; 1496 BP.
ID	AAQ28119	
XX	AAQ28119;	
AC	25-MAR-2003	(revised)
XX	08-FEB-1993	(first entry)
DT	25-MAR-2003	(revised)
DT	08-FEB-1993	(first entry)
XX	P.falciparum	LSA gene 3' region.
DE	P.falciparum	LSA gene 3' region.
XX		
KM	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope; paludism; liver stage-specific antigen; ss.	
OS	Plasmodium falciparum.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..1482
FT		/*tag= C
FT	repeat_region	/product= "LSA_C-terminal_region"
FT		37..639
FT	repeat_unit	/*tag= a
FT		37..87
FT		/*tag= b
FT		/rpe_type= TANDEM
XX		
XX	WO9213884-A1.	
XX	20-AUG-1992.	
XX	05-FEB-1992;	92WO-FR000104.
XX	05-FEB-1992;	91FR-00001286.
XX	05-FEB-1991;	91FR-00001286.
XX		
XX	(INSP) INST PASTEUR.	
XX		
XX	Guerinmarchand C, Druilhe P;	
XX		
XX	WPI; 1992-299985/36.	
XX	P-PSDB; AAR26944.	
XX		
XX	Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for	

ABK28180	DNA trans
ABJ33697	Human imm
AAH33181	Base sequ
AAH33182	Base sequ
AAH33180	Compox vi
AAH33184	Base sequ
ADH89675	Oligonuc1
ABJ15588	Oligonuc1
ABJ15589	Oligonuc1
ADH89401	Oligonuc1
AAH46686	Tumour su
ACI64817	M. xanthu
ABQ20939	Oligonuc1
ABQ20938	Oligonuc1
ABQ20942	Oligonuc1
ABQ20943	Oligonuc1
ABQ39210	Oligonuc1
ABJ39211	Oligonuc1
ABJ33013	Human imm
ABJ33958	Human imm
ABJ10199	Haematopo
ABJ10053	Haematopo
ABQ54066	Oligonuc1
ABQ54067	Oligonuc1
ABJ34358	Human imm
ACI64894	M. xanthu

PT vaccination against, treatment of and diagnosis of malaria.

XX Claim 2; Fig 8-10; 81pp; French.

CC This nucleotide sequence is the 3' part of the P.falciparum liver-stage
 CC specific antigen (LSA) gene. It codes for a polypeptide sequence which
 CC carries a T cell epitope characteristic of a protein produced in
 CC hepatocytes infected with P.falciparum. The polypeptide can be used in
 CC the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
 CC correct PN field.)

XX Sequence 1496 BP; 716 A; 169 C; 300 G; 311 T; 0 U; 0 Other;

Query Match 100.0%; Score 1482; DB 2; Length 1496;

Best Local Similarity 100.0%; Pred. No. 8.5e-202;

Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAGAACACAAAGCGATCTAGAACAGAGACGCTGCTAAAGAAAGTTGCAGAACAA 60
DB 1 CAGAACACAAAGCGATCTAGAACAGAGACGCTGCTAAAGAAAGTTGCAGAACAA 60
QY 61 CAAAGCGATTTAGAACAGATGACTTGTCTAAAGAAAGTTGCAGAGCAGAAAGCAT 120
DB 61 CAAAGCGATTTAGAACAGATGACTTGTCTAAAGAAAGTTGCAGAGCAGAAAGCAT 120
QY 121 TTAGAACAGAGAGAGCTTGTCTAAAGAAAGTTGCAGAACAAAGCGATCTAGAACAA 180
DB 121 TTAGAACAGAGAGAGCTTGTCTAAAGAAAGTTGCAGAACAAAGCGATCTAGAACAA 180
QY 181 GAGAGACGCTGTAAAGAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGAGAGCT 240
DB 181 GAGAGACGCTGTAAAGAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGAGAGCT 240
QY 241 GCTAAAGAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGAGAGAGAGAGAGAG 300
DB 241 GCTAAAGAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGAGAGAGAGAGAGAG 300
QY 301 AAGTTACAAAGAGCAGAACAGATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 AAGTTACAAAGAGCAGAACAGATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 GAACACAAAGCGATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GAACACAAAGCGATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 AGCGATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 AGCGATTTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 GAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 GAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 AGACGCTGTAAAGAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGAGAGAGAG 600
DB 541 AGACGCTGTAAAGAAAGTTGCAGAACAAAGCGATTTAGAACAGAGAGAGAGAGAG 600
QY 601 AAAAGAAAGTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 AAAAGAAAGTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 AATTTAGAAAGAAAGAAAGAGATGAGATATTTAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 AATTTAGAAAGAAAGAAAGAGATGAGATATTTAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 GAATATACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 780
DB 721 GAATATACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 780
QY 781 TCTTCTTTACCTCAGAGACAAAGAGAGAGATAGTAGAGATTCAGAGAGAGATTCATTA 840
DB 781 TCTTCTTTACCTCAGAGACAAAGAGAGAGATAGTAGAGATTCAGAGAGAGATTCATTA 840

```

```

QY 841 GAAAAACAAATAGAGATCTATTACAAACAAATGTTGAGAGAGAGAGATATACATATA 900
DB 841 GAAAAACAAATAGAGATCTATTACAAACAAATGTTGAGAGAGAGAGATATACATATA 900
QY 901 GGACATCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 GGACATCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 GCTGACATACAAATTCATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 961 GCTGACATACAAATTCATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 ATAGATAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 ATAGATAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 GAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 GAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GATGATTTAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 GATGATTTAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 TATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1321 TATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 TTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1381 TTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 GAGTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1482
DB 1441 GAGTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1482

```

RESULT 2
 AA028115
 ID AA028115 standard; DNA; 950 BP.
 XX
 AC AA028115;
 DT 25-MAR-2003 (revised)
 DT 08-FEB-1993 (first entry)
 DE P. falciparum LSA-R-NR coding sequence.
 KW Malaria; hepatocyte; sporozoite; plasmid DG 536; T-cell epitope;
 KW paludism; liver stage-specific antigen; ss.
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT repeat_region 3..629
 FT repeat_unit 3..53
 FT repeat_tag= b
 FN repeat_tag= b
 PD 20-AUG-1992.
 PD 05-FEB-1992; 92WO-FR000104.

```

XX 05-FEB-1991; 91FR-00001286.
XX (INSP) INST PASTEUR.
XX Guertinmarchand C, Druilhe P;
XX WPI; 1992-299985/36.
XX DR P-PSDB; AAR26941.
XX
XX Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
XX vaccination against, treatment of and diagnosis of malaria.
XX
XX Claim 17, Fig 2; 81pp; French.
XX
CC A genomic DNA bank of P.falciparum EcoRI fragments prepared in lambda
CC gtl1 was used to transform E.coli. The expression library was screened
CC with human antisera against antigens of all stages of P. falciparum. The
CC library was reseeded with antibodies affinity-purified on a clone
CC which was able to recognise antibodies specific to the hepatic phase.
CC About 40 clones were detected which produced a characteristic LSA
CC epitope. The clone with the largest insert (950 bases) encoded LSA-R-NR
CC containing a 12-repeat region followed by a non-repeat region. Preferred
CC antigenic polypeptides of the invention are derived from the amino acid
CC sequence of LSA-R-NR. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 950 BP; 456 A; 127 C; 206 G; 161 T; 0 U; 0 Other;
Query Match 64.1%; Score 950; DB 2; Length 950;
Best Local Similarity 100.0%; Pred. No. 2.9e-126;
Matches 950; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 AAGGCGATTAGAACCAAGAGACGCTCTAAGAAAAGTTGCAAGAACCAAGCGATT 70
Db 1 AAGGCGATTAGAACCAAGAGACGCTCTAAGAAAAGTTGCAAGAACCAAGCGATT 60
QY 71 TAGAACCAAGATAGACTTGTCTAAGAAAAGTTACCAAGAGACGCAAGATTAGAACCAAG 130
Db 61 TAGAACCAAGATAGACTTGTCTAAGAAAAGTTACCAAGAGACGCAAGATTAGAACCAAG 120
QY 131 AGAGACTTGTCTAAGAAAAGTTGCAAGAACCAAGCGATCTAGAACCAAGAGACGCTG 190
Db 121 AGAGACTTGTCTAAGAAAAGTTGCAAGAACCAAGCGATCTAGAACCAAGAGACGCTG 180
QY 191 CTAAAGAAAAGTTGCAAGAACCAAGCGATTAGAACCAAGAGACGCTCTAAGAAA 250
Db 181 CTAAAGAAAAGTTGCAAGAACCAAGCGATTAGAACCAAGAGACGCTCTAAGAAA 240
QY 251 AGTTGCAAGAACCAAGCGATTAGAACCAAGATGACGTGCTAAGAAAAGTTAGCAAG 310
Db 241 AGTTGCAAGAACCAAGCGATTAGAACCAAGATGACGTGCTAAGAAAAGTTAGCAAG 300
QY 311 AGCAGCAAGCGATTAGAACCAAGAGACGCTGCTAAGAAAAGTTGCAAGAACCAAA 370
Db 301 AGCAGCAAGCGATTAGAACCAAGAGACGCTGCTAAGAAAAGTTGCAAGAACCAAA 360
QY 371 GCGATTAGAACCAAGAGACGCTGCTAAGAAAAGTTGCAAGAACCAAGCGATTAG 430
Db 361 GCGATTAGAACCAAGAGACGCTGCTAAGAAAAGTTGCAAGAACCAAGCGATTAG 420
QY 431 AACCAAGAGACGCTGCTAAGAAAAGTTGCAAGAACCAAGCGATTAGAACCAAGAA 490
Db 421 AACCAAGAGACGCTGCTAAGAAAAGTTGCAAGAACCAAGCGATTAGAACCAAGAA 480
QY 491 GACGTGCTAAGAAAAGTTGCAAGAACCAAGCGATTAGAACCAAGAGACGCTGCTA 550
Db 481 GACGTGCTAAGAAAAGTTGCAAGAACCAAGCGATTAGAACCAAGAGACGCTGCTA 540
QY 551 AAGAAAAGTTGCAAGAACCAAGCGATTAGAACCAAGAGACGCTGCTAAGAAAAGT 610
Db 541 AAGAAAAGTTGCAAGAACCAAGCGATTAGAACCAAGAGACGCTGCTAAGAAAAGT 600
QY 611 TGCAGAGACGCAAGAGATTAGAACCAAGAGACGCTGCTAAGAAAAGTTAGAAA 670

```

```

Db 601 TGCAGAGACGCAAGAGATTAGAACCAAGAGACGCTGCTAAGAAAAGTTAGAAA 660
QY 671 GAAAAAAGCAATGAGATATATTAGCAGAGATTATATGTCGTTAGAAAATACAG 730
Db 661 GAAAAAAGCAATGAGATATATTAGCAGAGATTATATGTCGTTAGAAAATACAG 720
QY 731 CTATGAATCTTCATCAGAAAATGAAACGTGATATATTATACCATCAATCTTCTTAC 790
Db 721 CTATGAATCTTCATCAGAAAATGAAACGTGATATATTATACCATCAATCTTCTTAC 780
QY 791 CTCAGAGCAACAGAGGAATAGATAGATTCCAGAGAAATATCTAATATAGAAAAACA 850
Db 781 CTCAGAGCAACAGAGGAATAGATAGATTCCAGAGAAATATCTAATATAGAAAAACA 840
QY 851 ATAGGAATCTATTACCAACAAATGTTGAAGACGAAGGATATACATAAGACATCTTG 910
Db 841 ATAGGAATCTATTACCAACAAATGTTGAAGACGAAGGATATACATAAGACATCTTG 900
QY 911 AAGAAAAGAAATGTTCTAATTAACCGAACCAAAAAAGAAATTAATCT 960
Db 901 AAGAAAAGAAATGTTCTAATTAACCGAACCAAAAAAGAAATTAATCT 950
RESULT 3
AA080916
ID AA080916 standard; cDNA, 1320 BP.
XX
XX AA080916;
XX AC
XX 16-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 25-AUG-1995 (first entry)
XX DE
XX Plasmodium falciparum LSA-1 repeatless gene cDNA.
XX
XX Plasmodium falciparum LSA-1 repeatless gene; recombinant poxvirus;
XX multicomponent multistage malarial vaccines; immunogens;
XX malarial diagnosis; 88.
XX
XX OS Plasmodium falciparum; (PLSARPLS.14L.1).
XX
XX PN WO9428930-A1.
XX
XX PD 22-DEC-1994.
XX PF 10-JUN-1994; 94WO-US006652.
XX
XX PR 11-JUN-1993; 93US-00075783.
XX PR 09-JUN-1994; 94US-00257073.
XX
XX PA (VIRO-) VIROGENETICS CORP.
XX
XX PI Paolletti E, De Taiene C, Tine JA;
XX PT WPI; 1995-036113/05.
XX PS
XX Claim 3; Fig 11; 183pp; English.
XX
XX AA080916 is the P. falciparum LSA-1 repeatless gene cDNA sequence. New
XX recombinant poxviruses containing either the SPRA, ABRA, Pfhp70, AMA-1,
XX Pf825, Pf816, CSP, PfSSP2, LSA-1, LSA-1 repeatless, MSA-1, MSA-1 (N-
XX terminal p83 or C-terminal gp42) genes, or a combination of these in non-
XX essential regions of their genomes are claimed. These poxviruses (pref.
XX with a virulence reducing genomic deletion or disruption) can be used as
XX vaccines against malaria and for the prodn. of Plasmodium immunogens.
XX These viruses provide multicomponent, multistage vaccines due to their
XX expression of sporozoite, liver stage, blood stage and sexual stage
XX proteins. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-

```

CC OCT-2003 to standardise OS field)

XX Sequence 1320 BP; 619 A; 128 C; 227 G; 346 T; 0 U; 0 Other;

Query Match 57.6%; Score 853.2; DB 2; Length 1320;
 Best Local Similarity 87.9%; Pred. No. 1.5e-112;
 Matches 930; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```

Oy 425 ATTGAAACAAAGAGAGAGCTTGTAAAGAAAGTTGCAAGACAAACGATTTAGAAC 484
Db 263 ATGTAAAAATGTGTACAAACAAATTTTAAAGAACTTGTGTGTTTCKG 322
Oy 485 AAGAGAGAGCTGTAAAGAAAGTTGCAAGAAACAAACGATTTAGAACAGAGAGAC 544
Db 323 AGAATATATTCCTTAAAGAAATTAATTAATAGAGAGAAATTAATTAACATCA 382
Oy 545 GTGCTTAAGAAAGTTGCAAGAAACAAACGATTTAGAACAGAGAGAGCTGCTAAG 604
Db 383 TAAATGATGATGACGATTAAGAAATATATTAAAGGCGACAAACGAAACAGAGAG 442
Oy 605 AAAAGTTGCAAGAGCAGCAAGAGATTTAGAACAAAGGAGCTGATAGAAAAAAAT 664
Db 443 ATCTTGAAGAAAAAGCGCGCCCTTAAAGAAACGAGAGAGCTGATAGAAAAAAAT 502
Oy 665 TAGAAAGAAAAAGAAACATGAGATATATTAGCAGAGATTTATATGCTTTAGAA 724
Db 503 TAGAAAGAAAAAGAAACATGAGATGATTTAGCAGAGATTTATATGCTTTAGAA 562
Oy 725 TACAGCTTAAAGCTTCCATCAGAAATGAAACGTGATTTATATATACATCAATCT 784
Db 563 TACAGCTTAAAGCTTCCATCAGAAATGAAACGTGATTTATATATACATCAATCT 622
Oy 785 CTTTACCTCAGAACACAGAGGAAATGAGAGATTTCAAGGAAATATCTATATAGAA 844
Db 623 CTTTACCTCAGAACACAGAGGAAATGAGAGATTTCAAGGAAATATCTATATAGAA 682
Oy 845 AAACAAATAGAGATCTATTACAAACAAATGTTGAAGAGCAAGGATATATACATAAG 904
Db 683 AAACAAATAGAGATCTATTACAAACAAATGTTGAAGAGCAAGGATATATACATAAG 742
Oy 905 ATCTTGAAGAAAAAGAAAGATGTTCAATTAACAGAACAAAGAAAGATTAATCTG 964
Db 743 ATCTTGAAGAAAAAGAAAGATGTTCAATTAACAGAACAAAGAAAGATTAATCTG 802
Oy 965 ACATCAAAATCATCATTAAGAGACAGTAATATTTCTGATGTTATGTTTCAATTA 1024
Db 803 ACATCAAAATCATCATTAAGAGACAGTAATATTTCTGATGTTATGTTTCAATTA 862
Oy 1025 GTAAGTATGAGATGAATTAAGTCTGAATATGACGATTCATTAATAGATGAAGAG 1084
Db 863 GTAAGTATGAGATGAATTAAGTCTGAATATGACGATTCATTAATAGATGAAGAGAG 922
Oy 1085 ATGATGAAGACTTAAACGAAATTTAAGCCTATTTGTGCAATATGCAATTTCCAAG 1144
Db 923 ATGATGAAGACTTAAACGAAATTTAAGCCTATTTGTGCAATATGCAATTTCCAAG 982
Oy 1145 AAAACATAGGAATTTATTAAGAACTAGAGATTTGTAGAGAAAAATGAAATTTAG 1204
Db 983 AAAACATAGGAATTTATTAAGAACTAGAGATTTGTAGAGAAAAATGAAATTTAG 1042
Oy 1205 ATTTAGATGAAGAAATAGAAAAATCATCGAAGAAATTAATCTGAGAAAAATTAAG 1264
Db 1043 ATTTAGATGAAGAAATAGAAAAATCATCGAAGAAATTAATCTGAGAAAAATTAAG 1102
Oy 1265 GAAAGAAATATGAAAAACAAAGATTAATATTTTAAACCAATGATTAAGTTGTAT 1324
Db 1103 GAAAGAAATATGAAAAACAAAGATTAATATTTTAAACCAATGATTAAGTTGTAT 1162
Oy 1325 ATGAGCATTTAAAAATATTAAGAAATGATTAAGAGAAATTAAGTTGTAT 1384
Db 1163 ATGAGCATTTAAAAATATTAAGAAATGATTAAGAGAAATTAAGTTGTAT 1222
Oy 1385 TAAAAATCATTTGTTTATATATTTGACGAGACAAATGAATTTTACAGATCGTGAG 1444

```

```

Db 1223 TAAATCATTTGTTTATATATTTGACGAGCAATGAATTTTACAGATCGTGAGT 1282
Oy 1445 TATCTGAAGATTAATTAATTTTATGAAACATTA 1482
Db 1283 TATCTGAAGATTAATTAATTTTATGAAACATTA 1320

```

RESULT 4

ID AA005140 standard; DNA; 1566 BP.

XX AA005140;

XX AC 25-MAR-2003 (revised)

XX DT 05-NOV-1990 (first entry)

XX DE Sequence encoding N-terminal of peptide antigen to malarial sporozite.

XX KM Malaria; sporozite; vaccine; exoerythrocytic parasites; tetanus toxoid.

XX OS Plasmodium falciparum.

XX Key Location/Qualifiers

XX FT CDS 169..1554

XX FT /*tag= a

XX PN MO9006130-A.

XX PD 14-JUN-1990.

XX PF 30-NOV-1988; 88US-00278234.

XX PR 30-NOV-1988; 88US-00278234.

XX PR 12-APR-1989; 89US-00337204.

XX PA (BIOM-) BIOMEDICAL RES INST.

XX PI Hollingdal MR;

XX DR WPI: 1990-209624/27.

XX PS P-PSDB; AAR05766.

XX PT Novel malarial sporozite antigenic protein - useful as vaccine against

XX PS sporozite(s) and exo-erythrocytic parasites.

XX PS Disclosure; Page ?; -pp; English.

XX CC Antigen, preferably linked at the C-terminal to a carrier such as tetanus

XX CC toxoid, may be used as a vaccine against the malarial sporozite. (Updated

XX CC on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 1566 BP; 745 A; 210 C; 330 G; 281 T; 0 U; 0 Other;

XX Query Match 39.6%; Score 586.4; DB 2; Length 1566;

XX Best Local Similarity 91.7%; Pred. No. 1.1e-74;

XX Matches 620; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

```

Oy 1 CAAGAACAAACGATCTTGAACAAAGAGAGAGCTGCTTAAGAAAGTTGCAAGAACAA 60
Db 745 CAAGAACAAACGATCTTGAACAAAGAGAGAGCTGCTTAAGAAAGTTGCAAGAACAA 804
Oy 61 CAAGCGATTTAGAACAAAGATGAGCTGCTTAAGAAAGTTGCAAGAGCAGAT 120
Db 805 CAAGCGATTTAGAACAAAGATGAGCTGCTTAAGAAAGTTGCAAGAGCAGAT 864
Oy 121 TTAGAACAGAGAGACTTGTAAAGAAAGTTGCAAGAACAAACGATCTAGAACAA 180
Db 865 TTAGAACAGAGATGATCTTGTAAAGAAAGTTGCAAGAACAAACGATTTAGAACAA 924
Oy 181 GAGAGAGCTGTCTAAAGAAAGTTGCAAGAACAAACGATTTAGAACAGAGAGCT 240
Db 925 GATGAGAGCTGTCTAAAGAAAGTTGCAAGAACAAACGATTTAGAACAGATGAGCTT 984

```

QY 241 GCTAAAGAAAAGTTGCAAGAAACAAGCGATTGAAACAAGTACTGCTAAAGAA 300
 DB 985 GCTAAAGAAAAGTTGCAAGAAACAAGCGATTGAAACAAGTACTGCTAAAGAA 1044
 QY 301 AAGTTACAAAGACGCAAGAAAGCGATTGAAACAAGAGAGAGCGTCTAAAGAAAGTTGCA 360
 DB 1045 ACGTTGCAAGAGCGCAAGAAAGCGATTGAAACAAGAGAGAGCGTCTAAAGAAAGTTGCA 1104
 QY 361 GAACAACAAGCGATTGAAACAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAAACA 420
 DB 1105 GAACAACAAGCGATTGAAACAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAAACA 1164
 QY 421 ACGATTGAAACAAGAGAGCGTCTAAAGAAAGTTGCAAGAAACAAGCGATTG 480
 DB 1165 AAGATTGAAACAAGAGAGCGTCTAAAGAAAGTTGCAAGAAACAAGCGATTG 1224
 QY 481 GAACAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAAACAAGCGATTGAAACAAGAG 540
 DB 1225 GAACAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAAACAAGCGATTGAAACAAGAG 1284
 QY 541 AAGCGTCTAAAGAAAGTTGCAAGAAACAAGCGATTGAAACAAGAGAGCGTCT 600
 DB 1285 AAGCGTCTAAAGAAAGTTGCAAGAAACAAGCGATTGAAACAAGAGAGCGTCT 1344
 QY 601 AAGAAAGAGTTGCAAGAGAGAGCGATTGAAACAAGAGAGCGTCTAAAGAA 660
 DB 1345 AAGAAAGAGTTGCAAGAGAGAGCGATTGAAACAAGAGAGCGTCTAAAGAA 1404
 QY 661 AATTGAAAGAAAGAA 676
 DB 1405 TTGCAAGAAACAAGAA 1420

RESULT 5
 ADO21941
 ID ADO21941 standard; DNA; 1371 BP.

AC ADO21941;

DT 12-AUG-2004 (first entry)

DE LSA-NRC(H) construct DNA derived from Malaria parasite LSA-1.

KM LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;

KW parasite; sporozoite; vaccine; immunotherapy;

KM malaria parasite P. falciparum; LSA-NRC(H) construct; ds; gene.

OS Plasmodium falciparum.

XX Synthetic.

FT Key Location/Qualifiers
 FT CDS 1..1371
 FT /tag= a
 FT /product= "LSA-NRC(H) construct protein derived from
 FT Malaria parasite LSA-1"

PN WO200404167-A2.

XX 27-MAY-2004.

PD 12-NOV-2003; 2003WO-US036011.

PR 12-NOV-2002; 2002US-0425719P.

PA (REED-) REED ARMY INST RES WALTER.

PI Lander DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;

XX MPI; 2004-420309/39.

DR P-PSDB; ADO21942.

XX

PT Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
 PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
 PT (LSA-1) epitope.
 PS Claim 16; SEQ ID NO 25; 90pp; English.
 XX
 CC The invention relates to a novel recombinant LSA-NRC polypeptide
 CC comprising liver antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
 CC protein is found within the parasitophorous vacuole (PV) of Plasmodium
 CC falciparum. The polypeptide of the invention demonstrates antimalarial
 CC activities and may be useful as a vaccine during immunotherapy and as a
 CC diagnostic agent, preferably for diagnosing malaria. The current sequence
 CC is that of the LSA-NRC(H) Mut construct DNA of the invention. The
 CC construct is harmonised for expression in Escherichia coli and comprises
 CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
 CC parasite P. falciparum LSA-1 in addition to a C-terminal His6 tag.
 XX
 SQ Sequence 1371 BP; 539 A; 358 C; 283 G; 191 T; 0 U; 0 Other;

Query Match 34.0%; Score 503.4; DB 12; Length 1371;
 Best Local Similarity 63.1%; Pred. No. 6; 9e-63;
 Matches 791; Conservative 0; Mismatches 461; Indels 1; Gaps 1;

QY 227 AACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAAACAAGCGATTGAAACAAGATA 286
 DB 79 AACAGAGAAAGCATGAGAAAGAAACATGTGCTGAGCCACACTCTAGAGAAAGCTAA 138
 QY 287 GACTTCTAAAGAAAGTTGCAAGAGAGCAAGCGATTGAAACAAGAGAGAGCGTCTA 346
 DB 139 AACAGAGAAACAAACAATTTCTTGACAAAGAGCAAGAGCGTAGATGCAAGCTTAA 198
 QY 347 AAGAAAGTTGCAAGAGAGAGCGATTGAAACAAGAGAGAGCGTCTAAAGAA-2AG 405
 DB 199 AAGGATCCAGAGCCACTTTAAATCCCTCTGCGCAACCTCGGGGTTTCCAGAAACATC 258
 QY 406 TTGCAAGAAACAAGCGATTGAAACAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAA 465
 DB 259 TTTCTCAAAGAAACAAGCTGAAACAAGAGAGCAAGCTGATTGAAACATATCTACAGAC 318
 QY 466 CAACAAGCGATTGAAACAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAAACAAGAGC 525
 DB 319 GACGATGACAAAGAAATTAATTAAGGCCAGATGAAATCGCAGAGAGAGCTCGAA 378
 QY 526 GATTGAAACAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAAACAAGAGCGATTGAA 585
 DB 379 GAAAGAGCTGCTGAAACAGAGAGAGCGTCTGAAACAGAGCGCTGCTAAAGAAAGCTC 438
 QY 586 CAAGAGAGAGCGTCTAAAGAAAGTTGCAAGAGAGAGCGTCTGAAACAAGAGAG 645
 DB 439 CAGAGAGCGCTGCTAAAGAAAGCTCCAGAGAGAGAGCGGAGCTGAGAGAGCGCAAG 498
 QY 646 GCTGATAGCAAGAAAGTTGAAAGAAAGAAAGAGAACTGAGATATATTAAGAGAGAT 705
 DB 499 GCTGACAGAAAGAAAGAAAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
 QY 706 TTATATGCTGCTTGAAGAAATACAGCTATAGAACTTCAATGAGAAAGAGAGAGAT 765
 DB 559 CTGTACGCGCGCTGGAAGATCCAGCTATCGAACTCCATCGAAGAAAGAGAGAGAG 618
 QY 766 TATATAGCAATCAATCTTTTACCTAGAGAGAGAGAGAGAGAGAGAGAGAGATTCAG 825
 DB 619 TACATCCACACAG 678
 QY 826 GAAATATCTATATAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 885
 DB 679 GAAATCAGATCATGAG 738
 QY 886 AGGATATATCATTAAG 945
 DB 739 CGCGACATCCACAAAG 798
 QY 946 AAGAGAGATTAATGCTGAGATTCAGAAATATCATTAAGAGAGAGAGAGAGAGATTCAT 1005

Db 799 AAGAGAACAAAGCGCTGATATCCAGAAACAACCCCTGGAGACCGTGAACATTAGCGAC 858
QY 1006 GTTAATGATTTTCAATAAGTATGATGAGTGAATAAGTGAATATGACGATTCA 1065
Db 859 GTGAACGACTTCCAGATCGCAAGTACGAGACGAATATCCGCTGAAATGCAATACCTC 918
QY 1066 TTAATGATGAAGAAAGATGATGAAGACTTGAAGCAATTTTAAAGCTTATGCAATAT 1125
Db 919 CTGATCGACGAAGAAAGAACGACGAAGATCTGATGAATTCAAACCAATGCTCCAGTAC 978
QY 1126 GACAAATTCAGATGAGAAACATAGCAATTTTAAAGAACTGAAGATTGTAAG 1185
Db 979 GATTAATCTTCAGACGAAAGAAATATCGCATTTTCAAGAACTCGAAGACCTCATCGAG 1038
QY 1186 AAAAAATGAATAATTTGATGATTTAGATGAAGAAATGAAATTCATCGAAGATTATCT 1245
Db 1039 AAAAAAGAAACCTGACGACCTGACGAGGACATCGAAAAATCCTCGAAGAACTGAGC 1098
QY 1246 GAAAGAAAAATTAATAAGAAAGAAATATGAAAAACAAAGATATATTTTAAACCA 1305
Db 1099 GAAAGAAAAATCAAAAAAGCAAAATTCGAAAAAACCAAGACAACTTCAACCA 1158
QY 1306 AATGATTAATAAGTTTATGATGATGATATTAATAATAATGATTAAGCAGTTAT 1365
Db 1159 AACGCAATCCCTCTACGACGACGACATTTAAATAATCAAAAAACGACAAAGTGAAC 1218
QY 1366 AAGAAAAAGAAAAATTCATAAATCATTTGTTTCATATTTTGAAGGACAAATGAAT 1425
Db 1219 AAGAAAAAGAAAAATTTATCAAAATCCCTTTCACATCTTGAAGGCAATTAACGAAT 1278
QY 1426 TTACGATCGTGAATGATTTATCTGAAGATATATTAATAATTTTAAAGAACT 1478
Db 1279 CTGCAATTTGATGACGAAGTGAAGCAATTAATAATCTTATGTAAGCT 1331

RESULT 6
AD021919
ID AD021919 standard; DNA; 1374 BP.
XX
AC ADO21919;
XX
DT 12-AUG-2004 (first entry)
XX
DE LSA-NRC(H) Mut construct DNA derived from Malaria parasite LSA-1.
XX
XX LSA-NRC; liver stage antigen; LSA-1; T-cell; B-cell epitope;
XX parasitophorous vacuole; antimalarial; vaccine; immunotherapy;
XX malaria parasite P. falciparum; LSA-NRC(H) Mut construct; ds; gene;
XX mutant.
XX Plasmidium falciparum.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 1. 1374
FT CDS /*tag= a
FT /product= "LSA-NRC(H) Mut construct protein derived from
FT Malaria parasite LSA-1"
XX
XX WO200404167-A2.
XX
XX PD 27-MAY-2004.
XX
XX PF 12-NOV-2003; 2003WO-US036011.
XX PR 12-NOV-2002; 2002US-0425719P.
XX
XX (REED-) REED ARMY INST RES WALTER.
XX
XX Lanar DE, Hillier CJ, Lyon JA, Angov E, Kumar S, Rogers W;
PI Batdosa A;
XX
XX WPI; 2004-420309/39.

DR P-PSDB; AD021920.
XX
XX Recombinant LSA-NRC polypeptide for use as a vaccine and a diagnostic
PT agent, preferably for diagnosing malaria, comprises a liver stage antigen
PT (LSA-1) epitope.
XX
XX
XX
PS Claim 17; SEQ ID NO 3; 90pp; English.
XX
XX The invention relates to a novel recombinant LSA-NRC polypeptide
CC comprising liver stage antigen (LSA-1) T-cell and B-cell epitopes. LSA-1
CC protein is found within the parasitophorous vacuole (PV) of Plasmidium
CC falciparum. The polypeptide of the invention demonstrates antimalarial
CC activities and may be useful as a vaccine during immunotherapy and as a
CC diagnostic agent, preferably for diagnosing malaria. The current sequence
CC is that of the LSA-NRC(H) Mut construct DNA of the invention. The
CC construct is harmonised for expression in Escherichia coli and comprises
CC the harmonised N-terminal, C-terminal and 2 tandem repeats of the Malaria
CC parasite P. falciparum LSA-1 in addition to an amino acid insertion
CC within the T5 P. falciparum epitope and a C-terminal His6 tag.
XX
XX
SQ Sequence 1374 BP; 539 A; 358 C; 285 G; 192 T; 0 U; 0 Other;

Query Match 33.3%; Score 493.6; DB 12; Length 1374;
Best Local Similarity 63.1%; Pred. No. 1.7e-61;
Matches 793; Conservative 0; Mismatches 459; Indels 4; Gaps 2;

QY 227 AACAAAGAGACGCGCTTAAAGAAAGTTGCAAGAACAAAGGATTTAGAACAGATA 286
Db 79 AACGAGAAAGCATGAGAAAGAAACATGCTGTGAGCCCAACCTTACGAAAGACTTAA 138
QY 287 GACTTGTAAAGAAAGTTACAGAGACGCAAGCGATTTAGAACAGAGACGCTGTA 346
Db 139 AACAAAGAAACAAATTTTGAACAAGACAAAGAGCTGAGATGACAGCTTAA 198
QY 347 AAGAAAGTTCAGAACACAAAGCGATTTAGAACAGAGACGCTGTAAGA-AAG 405
Db 199 AACGATCCGACCACTTTAAATCCCTCTGCGCAACCTCGCGTTCGAGAACATC 258
QY 406 TTGCAAGAACAAAGCGATTTAGAACAGAGACGCTGTAAGAAAGTTGCAAGAA 465
Db 259 TTTCTCAAGAAACAACTGAACAGAGAGGCAAACTGATTAACATATCAACAGAC 318
QY 466 CAACAAAGCGATTTAGAACAGAGACGCTGTAAGAAAGTTGCAAGAACAAAGC 525
Db 319 GACGATGACAAAGAAATATCATTAAGGCCAGATGAAATTCGACGAAGACCTGAA 378
QY 526 GATTTAGAACAGAGACGCTGTAAGAAAGTTGCAAGAACAAAGCGATTTAGAA 585
Db 379 GAAAAAGCTGCTGAACAGCACTCGACCTGGAACAGGCGCCTCGCTAAAGAAAGCTC 438
QY 586 CAAGAGAGAGTGTAAAGAAAGTTGCAAGAGCAAGAAAGATTTAGAACAAAGAG 645
Db 439 CAGAGCGCTCGCTAAAGAAAGCTCGAGAGCAACGCGACCTGGAACAGCGCAG 498
QY 646 GCTGATACGAAAAAATTTAGAAAGAAAGAAAGCACTGAGATATTTAGCAGAGAT 705
Db 499 GCTGACGAGAAAAAAGCTGGAAGCAAGAAAGAAAGCAACGCGACGCTTCTGCTGAGAC 558
QY 706 TTAATATGCTGTTTAAATATCCAGCTATAGAACTTCATCAGAAATGAAAGTGAAT 765
Db 559 CTGTACGCGCGCTGGAATCCAGCTATCGAACTCCATCCGAAACGAACGCGGCTAC 618
QY 766 TATATACGACATCAATCTTTTACCTCAGAGCAACAGAGGGAATATGATGAGATTCCAG 825
Db 619 TACATCCACACAGAGACGCTTCCACAGATATATGCGGGAACCTCCGCGACGTAAG 678
QY 826 GAAATATCTAATATGAAAAAACAATAGAGAAATCTATTACAGAAATGTTGAGACAGA 885
Db 679 GAAATGACATCATCGAAAAACCAACGCGAAGCAATTAACCAACAGTGAAGCGCGC 738
QY 886 AGGATATATCAATAAGACATCTTGAAGAAAGAAAGATGTTCAATTAACCAAGACAA 945
Db 739 CGCGACATCCACAAAGGCCACTCGAAGAAAGAAAGCGGCTCCATCAACCAAGAACAG 798

Qy	946	AAAGAAGTAAATCTGTCGACATCAAAANTATATCATTTAGAGACAGTAAATATTTCGAT	1005
Db	799	AAAGAAGCAAAAGCGCTTGATATCCAGAACCAACCTCGAGAACGTAACATTAGCGAC	858
Qy	1006	GTTAATGATTTTCAAAATAGTAAGTATGAGATGAATAAAGTGCTGAATATGACGATCCA	1065
Db	859	GTGAACGACTTCCAGATCAGCAATGACAGAGACGAAATCTCCGCTGAATATGATGATCCC	918
Qy	1066	TTAAATAGATGAAGAAGATGATGAGACTTGAACGATTTAAAGCTATTGTGCAATAT	1125
Db	919	CTGATCGACGAAGAAGAGACGACGAATGATCGATGAATTTCAACCAATTTGTCOAATAC	978
Qy	1126	GACATATTTCCAAATGAGAAAGAAACCTAGGAATTTATAAAGACTTGAAGATTTGATGAG	1185
Db	979	GATTAATTTCCAGACGAAGAAGAAATATCGCACTTTACAAAGAACTCGAAGACCTATCGAG	1038
Qy	1186	AAAAATGAAATTTTAGATGATTTTGTAGTAGAAGAAATAGAAAAATCATCGAAGATTTATCT	1245
Db	1039	AAAAACGAAACCTGAGCAGCTGAGCGAAGGCATCGAAAAATCCTCGAAGAACTGAGC	1098
Qy	1246	GAAGAAAAATTAATAAAAAAGAAAGAAATATGAAAAAACAAAGATTAATATTTTAAACCA	1305
Db	1099	GAAGAAAAATTAATAAAAAAGCAAGAAATATCGAAAAAACCAAGACAAATCTTAACCA	1158
Qy	1306	AATGATAAAAATTGTATGATGAGCATTTAAAAAATTAATAAATGATTAACAGATTAT	1365
Db	1159	AACGCAAAATCCCTCTACGACGAGCACTTAAAAAATTCAAAAAGCAACCAAGTGAAC	1218
Qy	1366	AAGGAAAAAGAAAAATTCATTAATAATCATTTGTTCTATATATTTGACGAGACATGAAATT	1425
Db	1219	AAGGAAAAAGAAAAATTTATCAAAATCCCTCTCCACATCTTCGATGCGAATTAAGAAATT	1278
Qy	1426	TTACAGATCGGAGTGA---GTTATCGAAGATATTAATAATATTTATGAAACT	1478
Db	1279	CTGCAAAATTTAGACGAACGGTTGAGCGAAGACATCACTAAATCTTCATCAACT	1334

RESULT 7	AAQ28117	AAQ28117 standard; DNA; 988 BP.
XX	AC	AAQ28117,
XX	DT	25-MAR-2003 (revised)
XX	DT	08-FEB-1993 (first entry)
XX	DE	P.falciparum LSA gene 5' region.
XX	KW	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;
XX	KW	paludism; liver stage-specific antigen; ss.
XX	OS	Plasmodium falciparum.
XX	XX	
PH	Key	Location/Qualifiers
FT	5'UTR	1..32
FT		/*tag= a
FT	misc_feature	33..988
FT		/*tag= d
FT		/note= "5' part of LSA gene"
FT	repeat_region	492..988
FT		/*tag= b
FT	repeat_unit	492..542
FT		/*tag= c
XX		
PN	MO9213884-A1.	
XX		
PD	20-AUG-1992.	
XX		
PP	05-FEB-1992;	92MO-FR000104.
XX		
PR	05-FEB-1991;	91FR-00001286.
XX		

PA	(INSP) INST PASTEUR.
XX	
PI	Guerinmarchand C, Drulhe P;
XX	
DR	WPI; 1992-299985/36.
DR	P-PSDB; AAR26943.
XX	
PT	Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for
PT	vaccination against, treatment of and diagnosis of malaria.
XX	
PS	Claim 11; Fig 6; 81pp; French.
XX	
CC	This nucleotide sequence is the 5' part of the P.falciparum liver-stage
CC	specific antigen (ISA) gene. It codes for a polypeptide sequence which
CC	carries a T cell epitope characteristic of a protein produced in
CC	hepatocytes infected with P.falciparum. The polypeptide can be used in
CC	the preparation of vaccines against malaria. (Updated on 25-MAR-2003 to
CC	correct PN field.)
XX	
Q	Sequence 988 BP; 465 A; 120 C; 184 G; 219 T; 0 U; 0 Other;

Query Match	31.9%;	Score 472.8;	DB 2;	Length 988;
Best Local Similarity	95.7%;	Pred. No. 1.5e-58;		
Matches 486;	Conservative 0;	Mismatches 22;	Indels 0;	Gaps 0;

Qy	77	AAGTATACCTGCTAAAGAAAGGTTACAAAGCAGCAACGATTATTGAAACAAGAGAC	136
Db	481	AAGAAAAAGCAGCTAAAGAAAAGTTACAGGGGCAACAACGATTTCGAACAAAGAGAC	540
Qy	137	TTGCTAAAGAAAAGTTCGAAGAACAAACGCGATTGAAACAAGAGACGTGCTAAAG	196
Db	541	GTGCTAAAGAAAAGTTCGAAGAACAAACGCGATTGAAACAAGAGACGTGCTAAAG	600
Qy	197	AAAAGTTCGAAGAACAAACGCGATTGAAACAAGAGACGCGCTAAAGAAAAGTTGC	256
Db	601	AAAAGTTCGAAGAACAAACGCGATTGAAACAAGAGACGCGCTAAAGAAAAGTTGC	660
Qy	257	AAGAACAAACAAGCGATTTAGAACAAAGATAGACTTGCTAAAGAAAAGTTACAAGAGCAGC	316
Db	661	AAGAACAAACAAGCGATTTAGAACAAAGAGACCTTGCTAAAGAAAAGTTCAAGAACAC	720
Qy	317	AAAGCGATTTAGAACAAAGAGAGACGTGCTAAAGAAAAGTTCGAAGAACAAACGCGATT	376
Db	721	AAAGCGATTTAGAACAAAGAGAGAGCGTCTAAAGAAAAGTTCGAAGAACAAACGCGATT	780
Qy	377	TAGAACAAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACACAG	436
Db	781	TAGAACAAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAG	840
Qy	437	AGAGACTTGTCTAAAGAAAAGTTCCAGAACAAACAAGCGCTTTAGAACAAAGAGACGCTG	496
Db	841	AGAGACTTGTCTAAAGAAAAGTTCAAGAGACGAACGCGCTTTAGAACAAAGATAGACTTG	900
Qy	497	CTAAAGAAAAGTTCGAAGAACAAACGCGATTGAAACAAGAGAGACGTGCTAAAGAAA	556
Db	901	CTAAAGAAAAGTTCGAAGAACAAACGCGATTGAAACAAGAGAGACGTGCTAAAGAAA	960
Qy	557	AGTTGCAAGAACAAACGCGATTTAGA	584
Db	961	GGTTGCAAGAACAAACGCGATTTAGA	988

	RESULT 8
ADP85917	
ID	ADP85917 standard; DNA; 1300 BP.
XX	
AC	ADP85917;
XX	
DT	26-AUG-2004 (first entry)
KX	
DE	Synthetic construct #1.
XX	
Nanopore data analysis ; polymer ; polynucleotide;	

KW single-molecule resolution; ds.
 XX
 OS Synthetic.
 XX
 PN US2004110205-A1.
 XX
 PD 10-JUN-2004.
 XX
 PP 23-SEP-2003; 2003US-00668749.
 XX
 PR 23-SEP-2002; 2002US-0412959P.
 XX
 PA (WANG/) WANG H.
 XX
 PI Wang H;
 XX
 DR WPI; 2004-467656/44.
 XX
 PT Nanopore analysis system for analyzing polynucleotides includes nanopore
 PT device having structure with aperture, and nanopore data analysis system
 PT that can generate nanopore data points corresponding to each target and
 PT non-target polymer.
 XX
 PS Disclosure; SEQ ID NO 1; 17pp; English.

CC The invention relates to a nanopore analysis system that comprises a
 CC nanopore device and a nanopore data analysis system. The nanopore device
 CC has a structure having an aperture. The nanopore data analysis system can
 CC generate nanopore data points corresponding to each target polymer and
 CC each non-target polymer traversing the aperture, form a distribution
 CC pattern of the data points, and analyse a distribution of target polymer
 CC data points in the distribution pattern. The distribution pattern
 CC includes data cluster(s). The nanopore data analysis system is stored on
 CC a computer-readable medium. A mechanism is provided for analysing the
 CC distribution of target polynucleotide data points in the distribution
 CC pattern. The nanopore data analysis system can also analyse the
 CC distribution of the non-target polynucleotide data points. It can also
 CC determine a ratio between the target polynucleotide data points and the
 CC non-target polynucleotide data points. It can further analyse the
 CC distribution of target polynucleotide data points between the two data
 CC clusters; compare the distribution of the target polynucleotide data
 CC points between the two data clusters to a phosphorylation state standard
 CC distribution; and determine a ratio of phosphorylated target
 CC polynucleotide to non-phosphorylated target polynucleotides. It can also
 CC determine a cluster score for the target polynucleotide data points in a
 CC defined area, and compare the cluster score for the target polynucleotide
 CC data points to a cluster score for a chemical integrity standard density
 CC distribution for the defined area in a distribution of a target
 CC polynucleotide standard. The invention is useful for analysing polymers,
 CC such as polynucleotides. The invention nanopore analysis system
 CC potentially provides high speed sampling with single-molecule resolution,
 CC which may enable unprecedented dynamic range and sensitivity in analysis
 CC of samples containing charged polymers such as polynucleotides or
 CC polypeptides. The present sequence is a synthetic construct used in the
 CC invention.

Query Match	15.7%	Score 233.4	DB 12	Length 1300
Beet Local Similarity	48.7%	Pred. No. 1.5e-24		
Matches	633	Conservative	0	Matches 666; Indels 0; Gaps 0;
Qy	AAAGAAAGTTACAGAGCAGCAAGCGATTGAAACAAGAGAGCTGTTAAGAAAG	150		
Db	1 AA	60		
Qy	TTGCAGAGACAACAAGCGATTGAAACAAGAGAGCGCTTAAAGAAAGTTGCAGAA	210		
Db	61 AA	120		
Qy	CACCAAGCGATTGGAACAAGAGAGCGTGCTTAAAGAAAGTTCAAGAACACAAAGC	270		
Db	121 AA	180		

[illegible][illegible]


```

Db      2933 ATGAAGAAGAAAGAAACACATGAAAGAAAGAAATGTAACTATGAGAGAAAG 2992
Oy      1319 TGTATGATGACATATTAAAAATTAATAATGATAGACAGTTAATAGAGAAAGAAA 1378
Db      2993 AAGAAAAAGTACACATGAGAGAGAGAAAGTAAACATGAGAGAGAGAGAGAGAGAGTAA 3052
Oy      1379 AATTCAATAAATTCATTTGTTTCATATATTGACGAGACATGAAATTTTACAGATCGTG 1438
Db      3053 CACATGAGAGAGAGAGAGAGAGAGTAAATGATGAGAGAGAGAGAGAGAGAGAGAGAG 3112
Oy      1439 ATGAGTTATCTGAGATTAACCTAA 1463
Db      3113 AAGAGAGAGAGAGAGAGAGTAAAGAGAA 3137

RESULT 10
AAN80899
ID      AAN80899 standard; DNA; 210 BP.
XX
AC      AAN80899;
XX
DT      14-MAY-2003 (revised)
DT      25-MAR-2003 (revised)
DT      19-SEP-1990 (first entry)
XX
DE      Sequence encoding an epitope characteristic of a protein produced in
DE      liver cells infected with Plasmodium falciparum.
XX
KW      Plasmodium falciparum; vaccine; malaria diagnosis;
KW      malaria antibody detection.
XX
OS      Plasmodium falciparum.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..204
FT      /*tag= a
XX
PN      WO8805785-A.
XX
PD      11-AUG-1988.
XX
PP      09-FEB-1988; 88WO-FR000074.
XX
PR      09-FEB-1987; 87FR-00001543.
XX
PA      (INSP ) INST PASTEUR.
PA      (MARC/) MARCHAND C.
PA      (CNRS ) CENT NAT RECH SCI.
XX
PI      Drulilhe P, Pujjalome O, Marchand C, Langsley G;
XX
DR      WPI; 1988-235148/33.
DR      P-PSDB; AAP81355.
XX
PT      New polypeptide(s) contg. plasmodium falciparum epitope - useful for
PT      malaria vaccination or diagnosis.
XX
PS      Example 1; Fig 1; 37pp, French.
XX
CC      This sequence is contained in clone DG307 of a genomic DNA library of P.
CC      falciparum. AAN80899 is composed entirely of a repeating motif of 51 base
CC      pairs. The type of replication coded by DG307 seems to be conserved which
CC      implies that it may be present in all P. falciparum strains. The AA
CC      repeat unit is the basis of the peptides claimed in the patent, which are
CC      useful for the prodn. of malaria vaccines and as immunosassay reagents
CC      for diagnosis of malaria by detection of antibodies. (Updated on 25-MAR-
CC      2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
CC      (Updated on 14-MAY-2003 to correct PS field.)
XX
SQ      Sequence 210 BP; 97 A; 33 C; 47 G; 33 T; 0 U; 0 Other;
Query Match      12.4%; Score 183.2; DB 1; Length 210;

```

```

Best Local Similarity 95.9%; Pred. No. 2.2e-17;
Matches 188; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy      83 GACTCTGTAAGAAAGAGTTACAGACAGCAACCGATTGTGAAACAGAGAGACTTGCTA 142
Db      8 GACTTGCTTAAAGAAAGATTACAGAGAGCAAAACGATTGTGAAACAGAGAGACTTGCTA 67
Oy      143 AAGAAAGTTGCAAGAACCAAGCGATCTAGAACAGAGAGAGCGTCTAAAGAAAAGT 202
Db      68 AAGAAAGTTGCAAGAACCAAGCGATCTAGAACAGAGAGAGCTAGCTTAAAGAAAAGT 127
Oy      203 TGCAGAGAACCAAGCGATTTAGAACAGAGAGAGCGTCTAAAGAAAAGTTGCAAGAAC 262
Db      128 TACAGGGGCAACAAAGCGATCTAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 187
Oy      263 AACAAAGCGATTGTGA 278
Db      188 AACAAAGCGATTGTGA 203

RESULT 11
ADQ62833
ID      ADQ62833 standard; RNA; 1000 BP.
XX
AC      ADQ62833;
XX
DT      07-OCT-2004 (first entry)
XX
DE      Homopoly-A contaminant for RNaseH activity assay.
XX
KW      ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;
KW      fluorophore; fluorescence quencher; fluorescent signal; HIV;
KW      reverse transcriptase.
XX
OS      Synthetic.
XX
PN      WO2004059012-A1.
XX
PD      15-JUL-2004.
XX
PP      22-DEC-2003; 2003WO-US040879.
XX
PR      23-DEC-2002; 2002US-0436125P.
XX
PA      (AMHP ) WYETH.
XX
PI      Olson MW, O'Connell JF;
XX
DR      WPI; 2004-543471/52.
XX
PT      Detecting a nuclease-mediated cleavage of a target nucleic acid, useful
PT      for detecting and monitoring RNase H activity, comprises hybridizing a
PT      target nucleic acid to a fluorescently labeled oligonucleotide probe.
XX
PS      Example 1; SEQ ID NO 4; 61pp; English.
XX
CC      The invention relates to a method of detecting a nuclease-mediated
CC      cleavage of a target nucleic acid or measuring a RNaseH activity of an
CC      agent by hybridizing a target nucleic acid to a fluorescently labelled
CC      oligonucleotide probe complementary to the target nucleic acid and
CC      containing a fluorophore at one terminus and a quenching group at the
CC      other terminus and contacting the probe-target hybrid with an agent
CC      having nuclease activity. When the oligonucleotide probe is unhybridised
CC      to the target nucleic acid, the probe adopts a conformation that places
CC      the fluorophore and quencher in such proximity that the quencher quenches
CC      the fluorescent signal of the fluorophore. Formation of the probe-target
CC      hybrid causes sufficient separation of the fluorophore and quencher to
CC      reduce quenching of the fluorescent signal of the fluorophore followed by
CC      contacting the probe-target hybrid with an agent having nuclease activity
CC      to selectively cleave the target nucleic acid and thus release the intact
CC      probe. Detection of the release of the probe is by measuring a decrease
CC      in the fluorescent signal of the fluorophore as compared to the signal of
CC      the probe-target hybrid. The methods are useful for detecting and

```

CC monitoring RNaseH activity in real time. The methods are also useful for
 CC monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
 CC the invention, the activity of the RNaseH from E. coli or HIV reverse
 CC transcriptase is measure by the method of the invention. In order to test
 CC the specificity of the RNaseH cleavage of the target-probe complex, a DNA
 CC or RNA contaminant is added to the assay reaction. This sequence
 CC corresponds to the single stranded homopolymeric polyA RNA contaminant
 CC for testing this activity.

XX Sequence 1000 BP; 1000 A; 0 C; 0 G; 0 T; 0 U; 0 Other;

Query Match 12.2%; Score 180.2; DB 12; Length 1000;

Best Local Similarity 48.8%; Pred. No. 5.3e-17;

Matches 485; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

```

QY 397 AAAAAAAAAATTGCAGAACCAACGCGATTGAAACAAGAGAGACTTCTAAAGAAAG 456
   |||||
DB 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60
   |||||
QY 457 TTGCAGAACCAACGCGATTGAAACAAGAGAGCGCTTAAAGAAAGTTGCAGAA 516
   |||||
DB 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120
   |||||
QY 517 CAACAAGCGATTAGACAGAGAGCGCTTAAAGAAAGTTGCAGAACCAACAGC 576
   |||||
DB 121 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180
   |||||
QY 577 GATTGAAACAAGAGAGCGCTGTAAGAAAGTTGCAGAGCAGAAAGATTGAA 636
   |||||
DB 181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 240
   |||||
QY 637 CAAGAAGCGCTGATCGAATAAATTGAAAGAAAGAAACATGAGATATTTA 696
   |||||
DB 241 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300
   |||||
QY 697 GCAGAGATTATATGCTGTTAGAAATACGAGATATGAACTTCATCAGAAATGAA 756
   |||||
DB 301 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 360
   |||||
QY 757 CGTGATATTATATACACATCATCTCTTCTCAGAGACAGAGGAATAGTAA 816
   |||||
DB 361 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420
   |||||
QY 817 GATTCCAGGAATATCTATATAGAAACAAATAGAGATCTATTCAACAATGTT 876
   |||||
DB 421 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 480
   |||||
QY 877 GAAGAGCAAGGATATATACATAAAGACATCTTGAGAAAGAAAGATGTTCAATAAA 936
   |||||
DB 481 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 540
   |||||
QY 937 CGAAGAACAAAGAGATTAATCTGCTGACATACAAATCATACATTAGACAGTAAAT 996
   |||||
DB 541 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 600
   |||||
QY 997 ATTTCGATGTTAATGATTTTCAATTAAGTATGAGAGTAAATTAAGTCTGATAT 1056
   |||||
DB 601 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 660
   |||||
QY 1057 GACGATTCATTAATGATGAGAGAGAGATGATGAAGCTTGAAGCAATTTAAGCTATT 1116
   |||||
DB 661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 720
   |||||
QY 1117 GTGCAATATGACAAATTTCCAGATGAGAAACATAGAAATTTATAGACATAGAGAT 1176
   |||||
DB 721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 780
   |||||
QY 1177 TTGATAGAGAAATTTGAATTTAGTGTGTTTATGAGAAATGAAATCATCAGAA 1236
   |||||
DB 781 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 840
   |||||
QY 1237 GAATTAATCTGAGAGAAATTAAGAAAGAAATATGAAAGAAACAAAGATATATAT 1296
   |||||

```

```

DB 841 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900
QY 1297 TTTAAACCAATGATTAATGTTGATGATGACATATTTAAATAATGATTAAG 1356
   |||||
DB 901 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 960
   |||||
QY 1357 CAGGTTAATAGGAAAGAAATTCATTAATA 1389
   |||||
DB 961 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 993
   |||||

```

RESULT 12

AD062832/c

ID AD062832 standard; RNA; 1000 BP.

XX AD062832;

XX 07-OCT-2004 (first entry)

XX Homopoly-U contaminant for RNaseH activity assay.

XX ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;

XX fluorophore; fluorescence quencher; fluorescent signal; HIV;

XX reverse transcriptase.

XX Synthetic.

XX WO2004059012-A1.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040879.

XX 23-DEC-2002; 2002US-0436125P.

XX (AMDP) WYETH.

XX Olson MM, O'Connell JF;

XX WPI; 2004-543471/52.

XX Detecting a nuclease-mediated cleavage of a target nucleic acid, useful

XX for detecting and monitoring RNase H activity, comprises hybridizing a

XX target nucleic acid to a fluorescently labeled oligonucleotide probe.

XX Example 1; SEQ ID NO 3; 61pp; English.

XX The invention relates to a method of detecting a nuclease-mediated
 CC cleavage of a target nucleic acid or measuring a RNaseH activity of an
 CC agent by hybridizing a target nucleic acid to a fluorescently labelled
 CC oligonucleotide probe complementary to the target nucleic acid and
 CC containing a fluorophore at one terminus and a quenching group at the
 CC other terminus and contacting the probe-target hybrid with an agent
 CC having nuclease activity. When the oligonucleotide probe is unhybridised
 CC to the target nucleic acid, the probe adopts a conformation that places
 CC the fluorophore and quencher in such proximity that the quencher quenches
 CC the fluorescent signal of the fluorophore. Formation of the probe-target
 CC hybrid causes sufficient separation of the fluorophore and quencher to
 CC reduce quenching of the fluorescent signal of the fluorophore followed by
 CC contacting the probe-target hybrid with an agent having nuclease activity
 CC to selectively cleave the target nucleic acid and thus release the intact
 CC probe. Detection of the release of the probe is by measuring a decrease
 CC in the fluorescent signal of the fluorophore as compared to the signal of
 CC the probe-target hybrid. The methods are useful for detecting and
 CC monitoring RNaseH activity in real time. The methods are also useful for
 CC monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
 CC the invention, the activity of the RNaseH from E. coli or HIV reverse
 CC transcriptase is measure by the method of the invention. In order to test
 CC the specificity of the RNaseH cleavage of the target-probe complex, a DNA
 CC or RNA contaminant is added to the assay reaction. This sequence
 CC corresponds to the single stranded homopolymeric polyU RNA contaminant
 CC for testing this activity.

Sequence 1000 BP; 0 A; 0 C; 0 G; 0 T; 1000 U; 0 Other;

Query Match 12.2%; Score 180.2; DB 12; Length 1000;
Best Local Similarity 48.8%; Pred. No. 5.3e-17;
Matches 485; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

```

Qy 397 AAAGAAAGTTGCAAGAACAAACAAAGCGATTAGAACAGAGAGACTTGTAAGAAAG 456
Db 1000 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 941
Qy 457 TTGCAAGAACAAACGAGTTTAGACAGAGAGAGCTGCTAAAGAAAGTTGCAAGAA 516
Db 940 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 881
Qy 517 CAACAAAGCGATTAGAACAGAGAGAGCTGCTAAAGAAAGTTGCAAGAACAAAGC 576
Db 880 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821
Qy 577 GATTAGAACAGAGAGAGAGCTGCTAAAGAAAGTTGCAAGAGCAGCAAGAGATTAGAA 636
Db 820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
Qy 637 CAAGAGAGGCTGATCGAAGAAAAATTAGAGAAAAAGAAACATGAGATATTATTA 696
Db 760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
Qy 697 GCAGAGATTATATGTCGTTAGAAATPACAGCTATAGAACTTCATCGAAATGAA 756
Db 700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641
Qy 757 CGTGATATTATATACACATCAATCTTTACTCTGAGACAGAGAGGAAATAGTGA 816
Db 640 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 581
Qy 817 GATTCCAGAAATCTTATATAGAAAAACAATAGAACTCTATTACACAAATGTT 876
Db 580 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 521
Qy 877 GAAGAGAGAGGATATCATTAAGAGATCTTGAGAGAAAGAAAGATGTTCAATATAA 936
Db 520 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 461
Qy 937 CCAGAACAAAGAGATTAATCTGCTGACATACAAATTCATACATTAGACAGTAAT 996
Db 460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
Qy 997 ATTTCTGATGTTATGATTTTCAAAATAGTAAGATGAGATGAATAGTCTGAATAT 1056
Db 400 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341
Qy 1057 GACGATTGATTATGATGAAGAGAGATGATGAAGCTTAGACGATTTAGCTATT 1116
Db 340 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 281
Qy 1117 GTGCAATATGCAATTTCCAGATGAGAGAAACATAGCAATTTATTAAGAACTAGAGAT 1176
Db 280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221
Qy 1177 TTGATGAGAAATGAAAATTTAGATGATTAGTAGAGATAGAAAATCATCAGAA 1236
Db 220 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 161
Qy 1237 GAATTAATCGAAGAAAGAAATTAAGAAAGAAATATGAAAGAAAGATTAATAT 1296
Db 160 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101
Qy 1297 TTTAAACCAATGATAAAGTTTGTATGATGAGCATATTAAATAATATTAAGATATAG 1356
Db 100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 41
Qy 1357 CAGGTTAATAGGAAGAAAGAAATTCATATAA 1389
Db 40 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

```

RESULT 13

AAQ87587
ID AAQ87587 standard; DNA, 1686 BP.

AC AAQ87587;

DT 27-AUG-2003 (revised)

DT 19-DEC-1995 (first entry)

DE DNA encoding Leucocytozoan protozoa structural protein epitope.

KW leucocytozoan protozoa; structural protein; epitope; vaccine; fowl;

KW leucocytozoanosis; treatment; ss.

OS Leucocytozoan.

PN JP07089995-A.

PD 04-APR-1995.

XX 10-SEP-1993; 93JP-00226078.

XX 10-SEP-1993; 93JP-00226078.

XX (DOBU) DOBUTSUO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.

XX (NIRS-) NISSEIKEN KK.

XX WPI; 1995-167252/22.

XX P-PSDB; AAR70491.

PT Immune inducing polypeptide against Leucocytozoan protozoa - useful in

PT production of vaccines for treatment of leucocytozoanosis in fowl.

XX Claim 1; Page 12-14; 20pp; Japanese.

CC AAQ87587-89 encode polypeptides having a whole or partial epitope of a

CC structural protein of Leucocytozoan protozoa (see AAR70491-93). The

CC polypeptides and DNA encoding them are useful in the production of

CC vaccines for the treatment of leucocytozoanosis of fowl. (Updated on 27-

XX AUG-2003 to correct OS field.)

XX Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 U; 0 Other;

Query Match 12.1%; Score 179.4; DB 2; Length 1686;
Best Local Similarity 46.9%; Pred. No. 6.6e-17;
Matches 595; Conservative 0; Mismatches 671; Indels 3; Gaps 1;

```

Qy 2 AAGAACAAAGCGATCTAGAACAGAGAGAGCTGCTAAAGAAAGTTGCAAGAACAC 61
Db 101 AAGAGAAATGAAAGAAATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 160
Qy 62 AAAGCGATTAGAACAGATAGACTTGCTAAAGAAAGTTACAGAGAGCAAGCGATT 121
Db 161 AACAAGAAACAAGAAATGTAAGAACAGAAACAAGATGAAGAGAACAGAGAGAG 220
Qy 122 TAGAACAGAGAGCTTGCTAAAGAAAGTTGCAAGAACAAAGCGATCTAGAACAG 181
Db 221 ATGAAGAAAGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAG 280
Qy 182 AGAGACGTCCTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGCTG 241
Db 281 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
Qy 242 CTAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGATAGACTTGCTAAAGAA 301
Db 341 ATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
Qy 302 AGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Db 401 AAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 460

```

Qy	362	AAACAACAAAGCCGTTTAAAGACAAGAGACGTCGCTAAAGAAAAGTTGCCAAGAACCAACAA	421
Db	461	AAACAACAAAGAACAGATGAAAGAGAACCAAGATGAAGAAACAGATGAAGAAAG	520
Qy	422	GGCATTTAGAACAGAGAGACTTCTTAAGAAAAGTTGCCAAGAACCAAGAGCGATTAG	481
Db	521	AAAGAAAGAAACAGAGAAACAGATGAAAGAACAGAGAAAGATATCTGTAAGAAAGAA	580
Qy	482	AAACAAGAGACGTCGCTTAAGAAAAGTTGCCAAGAACCAAGAGCGATTAGAACAGAGA	541
Db	581	ATGAAGATGAAACAAAAAGAAAAGAAAGAAACAGAGATGAAAAATATATGTTG	640
Qy	542	GACGCGCTTAAGAAAAGTTGCCAAGAACCAAGCGATTAGAACCAAGAGACGTCCTA	601
Db	641	AAAAAGAAAAGATGAAGATATATGTAAGAAAAGAAACAGAACATTTTAAAGAAATTAG	700
Qy	602	AAAGAAAGTTGCAAGAGCAGCAAAAGAGATTAGAACAAAGAAAGCGCTGATACGAAAAAA	661
Db	701	AAAAAGAAAGTTTGAAGAAAAAGTGAAGAGTGAACA--TAAAGATGATTAAGAAAGACA	757
Qy	662	ATTAGAAAAGAAAAGAAACATGAGATATATTAGCAGAGATTTATATGTCGTTTAG	721
Db	758	CAAAATATAAAAAGCATGAAGTTGAAGAGAAACGAGAAATAGAAATTAAGAGGAATCAG	817
Qy	722	AAATTCACGACTTAGAACCTTCCATCAAGAAATGACGTCGATTTATATATCCATCAAT	781
Db	818	AAAGAAATGATAGAAAAGAAAATCAGAAAGTAAAGAACAAATGAGCAAGTCAAGAGAAAA	877
Qy	782	CTTCTTTACCTGAGGCAACAGAGGAAATAGTAGAATTCAGAGAAATTCATATATAG	841
Db	878	CAATATGAACAAAAGAGATATATACAGATCCAAATTTGGTCATCAATTTAAGATATATGA	937
Qy	842	AAAAAACAATAGAGAAATCTATTACAACAAATGTTGAAGACAGAGGATATATCATTAAG	901
Db	938	TAAATTAAGAAACAGAAAACAGAAAACCTATACGAAATTTATATCAATTTAAAGTAACTA	997
Qy	902	GACATCTTGAAGAAAAGAAAGATGTTCAATTAACCCAGAACCAAAAAGAAAGATTAATCTG	961
Db	998	ATGAATCAAAATCCGAAATTAATTCAACTCAAAATTAATTTGAAGATGAAAAATPAAAAATT	105
Qy	962	CTGACATACAAATCATATATAGAGACAGTAAATTTTCTGATGTTAATGATTTTCAA	102
Db	1058	CATCATTTGAAAAAACAGAAATTAATGAAGCTATCTAGAGGCGGAATTCGGAACAATAG	111
Qy	1022	TAACTAAGATGAGAGATGAATTAAGTCTGAATATGACGATTCATTAATAGATGAAGAG	108
Db	1118	TTGAAGAAAGTTCTTGAAGAAACACCGATGTTGAGAGAAAGATGTAGAGAAACACTG	117
Qy	1082	AAGATGATGAAGACTTAGACGATTTAAGCTATATGCTCAATATGACAAATTTCCAGATG	114
Db	1178	TAGTTGAAGAAAGAAATGATGAAGAAACACTGCTGATTTGAAGAAAGATGATGAAGAAA	123
Qy	1142	AAGAAAACATAGGAATTTTAAAGAACTAGAAATTTGATAGAGAAAATGAATAATTGAG	120
Db	1238	CACCGTAGTGAAGAGATGTCGTCGAAGAAACACCAATAGTGGAAAGAGATGATG	129
Qy	1202	ATGATTTAGATGAAGAAATGAAGAAATCATCGAAGAAATTAATCTGAAGAAAATTAATA	126
Db	1288	AAGAAACTTAATTAATGTTGAAGAAAGTATGTAAGAAACACCAATGATTTGAAGAAAGAAATGA	135
Qy	1262	AAGGAAGA 1270	
Db	1358	TAGAGAAA 1366	
RESULT 14			
AAA70099			
ID AAA70099 standard; DNA; 3579 BP.			
XX AAA70099;			
AC			
XT 07-NOV-2000 (first entry)			

D8			Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:232.
KW			Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX			antimalarial; malaria; protozoacide; infection; insecticide; ds.
OS			Plasmodium falciparum.
XX			
PX			WO200025728-A2.
DD			
PD			11-MAY-2000.
PF			05-NOV-1999; 99WO-USO26796.
PR			05-NOV-1998; 98US-0107131P.
PA	(HOFF//)	HOFFMAN S.	
PA	(CARU//)	CARUCI D.	
PA	(GARD//)	GARDNER M.	
PA	(VENT//)	VENTER J C.	
PI		Hoffman S,	Carucci D,
DX			Gardner M,
DR			Venter JC;
PT			WPJ: 2000-365347/31.
PT			
PS			Proteins encoded by chromosome 2 of the human malarial parasite, plasmodium falciparum, useful as antimalarial vaccines and in the diagnosis of P.falciparum infection. Disclosure: Page 457-458; 577pp; English.
CC			The present invention describes proteins and their fragments (I) encoded by chromosome 2 of the human malarial parasite, Plasmodium falciparum. Also described are: (1) nucleotide sequences (II) encoding (I); and (2) vaccines against P. falciparum infection comprising (I) or (II). (I) and (II) are useful for the development of vaccines against P. falciparum infection. (I) and polyclonal antisera or a monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with P. falciparum. Furthermore, (I) (especially when they are rifins or secreted or membrane proteins) can aid the identification of drugs to treat or prevent P. falciparum infection, or they can be used to identify drug resistance in P. falciparum. Sequencing of the Plasmodium chromosome 2 and the subsequent identification of proteins encoded by it will help to expand our understanding of parasite biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AAB70078 to AAB70287 and AAB18144 to AAB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification
SQ			Sequence 3579 BP; 1904 A; 398 C; 552 G; 725 T; 0 U; 0 Other;
Query Match			11.9%; Score 176.6; DB 3; Length 3579; Best Local Similarity 46.3%; Pred. No. 1.6e-16; Matches 661; Conservative 0; Mismatches 755; Indels 9; Gaps 2
OY	36	TGCTAAGAAAGAATTGCACGAACAACAAGCGATTTAGAACAGTAGACTTGCTTAAGA	95
DB	744	TGATTAATAACAAAAGAAAAAATACTACCAACCATGTGGAAAAAAAACAAAGAACAGAGA	803
OY	96	AAGGTTCACAGAGCAGCAACGCCATTTTGAACACAGAGAGACTTGTTAAGAAAAGTTGCA	155
DB	804	TAAAGAAACATTTCTAAAAAAAAGAAAAACAAAACAAATTTTAATTAAGAAAGAAAAACCA	863
OY	156	AGAACCAACAAAGCATCTAGAAACAAGAGAGCGTCTAAAGAAAAAGTTGCACAAACAACA	215
DB	864	AAATATAGAAAATCATATATCTCTCAAAAGTAATATTAACACAGACGAATGTGAAATTATGA	923
OY	216	AAGCGATTTAGACACAGAGAGACGTGCTAAAGAAAAGTTGCACAAACAACGCAATTT	275
DB	924	TGATTAATGAAAAAGAAATGGAAATCCCACAAAAATRTTAACAGAAAAAGGGGGGAATTGA	983

```

OY 276 AGAACAAATAGACTTGTCTAAAGAAAAGTTACAAGAGCAGCAAGCGATTGTGAACAAGA 335
DB 984 AAAAACTGATCTACTAATATATAGCAGATATATAAAAAAGAAAGAAAGAAACAAA 1043
OY 336 GAGAGCTGCTAAAGAAAAGTTGCAAGAACCAACAAACGATTTGAAACAAGAGAGCGTGC 395
DB 1044 AGATGAAAAAGAAAAAATTCACACAGCTAGTAAAGATGTACAATTAAATTAAGTTGG 1103
OY 396 TAAAGAAAAGTTGCAAGAACCAACGCGATTGAAACAAGAGACCTTGTCAAGAAA 455
DB 1104 AGAAGAAAAGCAAGAGCGAAGAAAAGAAAGCAAGAGGACAGATGTGAAGAAAGATAC 1163
OY 456 GTTGCAGAACCAACGCGATTGTAACAAGAGAGCGTGTAAAGAAAAGTTGCAAGA 515
DB 1164 AGATGATGAGAGAGTTACAGATGATGAGAGAGATACAGATGATGAAGAGATCAAGTGA 1223
OY 516 ACAACAAA-----CGCATTTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAGAAC 569
DB 1224 TGAAGAAAAGTACAGGTGATCAAGAAAAGCAAGAGAAACAGAAAGTACGAAAAAAGAAC 1283
OY 570 ACAAGCGATTTAGAACAGAGAGACGTGCTAAAGAAAAGTTGCAAGAGCGCAAGAGA 629
DB 1284 AGAAAAAGCCGAGAGAAAGTTAGAAAGAACCAAGAGAAATCAGAAAAAGCAAAAGAGA 1343
OY 630 TTTAGAACAAAGAGGCTGATACGAAAAAATTTAGAAAAGAAAAGAAACATGAGAGA 689
DB 1344 ATCCGAAAAAGTAAAGAAAGATCAAGAAAAGTAAAGAAAGATCAAGAAAAGCAAAAGA 1403
OY 690 TATATTAGCAGAGATTTATATATGTCGTTAGAAAATCCAGCTATAGAACTTCCATCAGA 749
DB 1404 AAAAACTGAGAGAAAGTGAAGAAAAAATCGAAGACGAAAAAGGAAAGGAAATACAAAAA 1463
OY 750 AAATGAGCGTGAATTTATATATCCATCAATCTTTACTCAGAGCAACAGAGGGA 809
DB 1464 AGAAACAGATGTAGTGAAGAAAAAAGGAGAAATATGAGAGGGAACAGATGATGA 1523
OY 810 TAGTGAAGATTCGAAGAAATATCTATATAGAAAAAACAATAGAGATCTATTACAC 869
DB 1524 AGAAGCAAAAGAAAAGAAAGAGAGAGAGAAACAAAGTAAAGAAAGAAAACAGA 1583
OY 870 AAATTTGAGAGCAAGAGGATATATCATTAAGAGCATCTTGAGAAAAGAAAGATGTTTC 929
DB 1584 AAAAGACAAAGAGGGAACAGATTATGAAGAGATACAGATGATCAGCAAAAGATGAAGA 1643
OY 930 AATAAACCAAGAAAGAAAGATTAATCTGCTGACATCAAAATCTACTTTGAAGAC 989
DB 1644 AACAAAGGTAGAAAGAAAAGAAAAGAGAGACGAAAGAACTGAAAGAGAGCAAAA 1703
OY 990 AGTAAATATTTCTGATGTTAATGATTTTCAATTAAGTAAATGAGATGAATTAAGTGC 1049
DB 1704 AGAAACAGAGTGAAGAAAAGAAAAGAAAAGAAAAGCAAGAGGGAACAGATTATGAAGA 1763
OY 1050 TGAATATGACGATTCATTAATGATGAAGAGAGATGATGAAGACTTATGACGAATTTAA 1109
DB 1764 AGATCAATGATGTACAGCAAAAGATGTAGAAAAGAAAGTAAAGAAACAGACGAGAGA 1823
OY 1110 GCCTTTGTGCAATGTGCAATTTCCAGATGAAGAAAACATAGCAATTTTAAAGAACT 1169
DB 1824 CAAGAGAAAAGCAAGAGGGAACAGATGATGAAGAAAGCAAGTGAAGAAACGACACT 1883
OY 1170 AGAAGATTTGATAGAGAAAATGAATAATTTAGATGATTAGTGAAGAAATGAAAATTC 1229
DB 1884 AGACGACCAAGAGAGAGCGAGAGAGATTAAGAGACGACAAAGAAAAGAGCAAAAGA 1943
OY 1230 ATCAGAGAGATTTATGTGAAGAAAATTAAGAAAAGAAAAGAAATGTGAAGAAAAGAGA 1289
DB 1944 AGATGACAAAGAGATGACAAAGAAAAGACAAAGAAAGTGAACAAAGAAAATACAAAGA 2003
OY 1290 TAAATATTTTAAACCAATGATTAAGAGTTTGTATGATGACATATTAAGAAATATTAAGA 1349
DB 2004 AGATGAC---AAAGAGATGACAAAGAAAGATGACAAAGAAAAGCAAGAGATTAAGAA 2060

```

```

OY 1350 TGATPAGCGGTTATATAGAAAAAGAAAATTCATTAATCATTTGTTATATTTGA 1409
DB 2061 AGAAAAAGACAAAGAGTTATCAAGAAAAAGACAAAGAGATGACAAAGAAAAGACAA 2120
OY 1410 CGAGACACATGAATTTTACAGATCGTGATGAGTTATCTGAAGATATA 1458
DB 2121 AGAAGATGACAAAGAAAAAGCAAAAGAAAGATTAACAAAGAAAAGACAAA 2169

RESULT 15
AAA70212
ID AAA70212 standard; DNA; 1998 BP.
AC AAA70212;
XX
XX 07-NOV-2000 (first entry)
DT
XX Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:345.
DE
XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozoacide; infection; insecticide; de.
OS Plasmodium falciparum.
PN WO200025728-A2.
XX
XX 11-MAY-2000.
PD
XX 05-NOV-1999; 99WO-US026796.
PF 05-NOV-1998; 98US-0107131P.
PR 05-NOV-1998;
XX
XX (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffmann S, Carucci D, Gardner M, Venter JC;
DR WPI; 2000-365347/31.
XX
XX Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PR diagnosis of P.falciparum infection.
XX
XX Disclosure: Page 537; 577pp; English.
XX
XX The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II); (I) and
CC (II) are useful for the development of vaccines against P. falciparum
CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
CC immunogens comprising the sequences of (I), are useful in the detection
CC of infection with P. falciparum. Furthermore, (I) (especially when they
CC are refined or secreted or membrane proteins) can aid the identification
CC of drugs to treat or prevent P. falciparum infection, or they can be used
CC to identify drug resistance in P. falciparum. Sequencing of the
CC Plasmodium chromosome 2 and the subsequent identification of proteins
CC encoded by it will help to expand our understanding of parasite biology,
CC a process hampered by the complexity of the parasitic life cycle, and
CC provide new targets for resistance and drug development. Parasite resistance
CC to drugs and mosquito resistance to insecticides have led to a resurgence
CC of malaria in many parts of the world, and there is a pressing need for
CC vaccines and new drugs. AAH70078 to AAA70287 and AAB18144 to AAB18352
CC represent nucleotide and protein sequences given in the present
CC invention, but which are not specifically mentioned within the
CC specification.
XX
SO Sequence 1998 BP; 1127 A; 141 C; 338 G; 392 T; 0 U; 0 Other;
Query Match 11.5%; Score 170.6; DB 3; Length 1998;
Best Local Similarity 45.7%; Pred. No. 1.2e-15;

```

Matches	594;	Conservative	0;	Mismatches	706;	Indels	0;	Gaps	0;
Qy	80	ATAGACTTCTAAAGAAAGTTTCAAGAGACGAAAGCGATTTTGAACAAGAGACTTG	139						
Db	413	ATAGACAAAATGATTTAAATTTTCAAAAGTGAAGAAAAATGACAAAGCATTAATTTAAAAAG	472						
Qy	140	CTTAAGAAAAGTTGCAAGAACAAAGCGATCTAGAACAGAGACGCGCTTAAAGAA	199						
Db	473	AAAAAGAAAACGAGATATCTCTAATTCATGCAAGAACAAAAGATGTTAAAGAAAG	532						
Qy	200	AGTTGCAGAACCAACAAAGCGATTTAGAACAGAGACGCTGTAAAGAAAAGTTGCAAG	259						
Db	533	GAGTAAAGAAATTTGAAGAAAAGAAAAGAAAATTTTCAGATGATCTAATAAGTGG	592						
Qy	260	AACACAAAGCCGATTTAGAACAGATAGACTTGCTTAAAGAAAAGTTTCAAGAGCAGCAA	319						
Db	593	AGGAAAACAAAATACTGATGACCATTAAGTTGAGAGAAAACAAAATACTGATGATCAAT	652						
Qy	320	GGGATTTAGAACAGAGACGCTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAG	379						
Db	653	AAAGTGAGGAAAACAAAATACTGATGATCTAATAATGAGAGAGTTAAAGAAAAGTAGAG	712						
Qy	380	AACAGAGAGACGCTGTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGA	439						
Db	713	AACATGAAGAGACGAAAGAGAGGATTAAGAAAAGAAAATTCAGAAAACAAAATAAG	772						
Qy	440	GACTTGCTTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCTGCTA	499						
Db	773	ATGAAAATTAAGTGAAGAAAATGATGAAGATTAAGTAAATTAAGTGAAGATGAAGTTG	832						
Qy	500	AAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGACGCTGTAAAGAAAAGT	559						
Db	833	ATGATGATTTGAGAGAGATTAATAATGAATAATGATGATTAATGATGATTAATAAGAAA	892						
Qy	560	TGCAAGAACCAAGACGATTTAGAACAGAGACGCTGTAAAGAAAAGTTGCAAGAGC	619						
Db	893	CAGATTAACACATTTTGAAGAAAGAGAAAATGAATAATTTGAAAAGCAATTTAGATGA	952						
Qy	620	ACCAAGAGATTTAGAACAAAGAGCGCTGATTCGAAAAAAATTTAGAAAAGAAAAGG	679						
Db	953	ACAAAAGATGTTAAATAATTAAGATATCAAGAACGAAAAAGTAAAGTACTGAGAAAG	1012						
Qy	680	AACATGAGATATATTAGACAGAGATTTATATGCTTGTAGAAAATACAGCTATTGAAC	739						
Db	1013	AAAAAAGTTAAAGATTTAGAAAGGAAAGAAAAGTTAAAGATTAAGAGAAAGAAAAAGTAAAG	1072						
Qy	740	TTCATCAGAAAATGAACTGTGATATATATATACCATCAATCTTCTTACTCAGAGACA	799						
Db	1073	ATTAAGAGAGAAAAGTTAAAGATTAAGAGAGAAAAGTAAAGATTTAGAGAAAG	1132						
Qy	800	ACAGAGGGAATGTAGAGATTCACAGGAAATCTTATTAATAGAAAAACAAATAGAGANT	859						
Db	1133	AAAAAGAAAAGATTAAGATTTAGAGAGAGAAAAGAAAAGTAAATCTCGAGAGAAAAAG	1192						
Qy	860	CTATTACACAAATGTTGAAGACGAGAGGATTAATCATTAAGAGACCTCTGAAGAAAAGA	919						
Db	1193	AAAAAGATTAAGATTTAGAGAAAAGAAAAGCAAGATATGAAAAATTTAAAGATTAAC	1252						
Qy	920	AAAGTGTCTCAATTAAGAACAGAACAAAAGATTAATCTCTGACATACAAAATCATATA	979						
Db	1253	AAAAATGATGAAGAAAAGAAAAGCGATTAAGAAAAGAAAAGAAAATGATTAACAGATATAC	1312						
Qy	980	CATTAGACAGCTAAATATTTCTGATGTTAATGATTTTCAAAATAGTAAGTATGAGATG	1039						
Db	1313	ATGATGATTAATGATGTAAGAAAATGATGAGAAATTAAGAAAATGATGTCGAAAGAG	1372						
Qy	1040	AAATAGTCTGAATATGACGATTCATTAATGATGAGAGAGAGATGATGAGACTTAG	1099						
Db	1373	ATGAGAGATGAAGCATGAGAAAACMAAAAAAAGAAAAAAGAAAAATGAGAAATGAAA	1432						
Qy	1100	ACGAAATTAAGCCTATTGTGCAATATGACAAATTTCCAAAGATGAGAAAACATAGCAATT	1159						
Db	1433	ATGGAATGAAAAATGGAAGTGAAGAAATGGAATGAAAATGAAAATGGAATGAAA	1492						

Qy	1160	ATTAAGAACTAGAAAGTTTGTATGAGAAAAATGAAAATTTAGATGTTTATGATGAAGCA	1219
Db	1493	ATGAAAATTTAAAAATGAAAGTGAATAATGAAAATGAAAATGAAAATGAAAATGAAA	1552
Qy	1220	TAGAAAAATCATCAGAAAGATTAATCTGAGAGAAAATTAATAAGAAAGAAATATGAAA	1279
Db	1553	ACGAAAATTAATAACGAAAAAGAAAACGAAAAAGATTAATAATTTAAAGAGATTGAAAAG	1612
Qy	1280	AAACAAAGATTAATAATTTTAAACCAATGATTAAGATTGTATGATGAGATATTTAA	1339
Db	1613	TAACAATGCAACAAGAAAATCTATGAAAAATTAATTAATAATCTGAAATTAACATTA	1672
Qy	1340	AATTAATAATGATTAAGACGTTTATTAAGAAAAGAAA	1379
Db	1673	CAAAATCAAAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1712
RESULT 16			
AAS46735/C			
ID	AAS46735 standard; DNA; 6292 BP.		
XX	AAS46735;		
XX	18-DEC-2001 (first entry)		
DT			
XX	Tumour suppressor gene derived chemically modified sequence #459.		
DE			
XX	Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;		
KW	tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;		
KW	cytosine methylation; de.		
OS	Homo sapiens.		
XX	WO200168912-A2.		
XX	20-SEP-2001.		
XX	15-MAR-2001; 2001WO-BP002955.		
PF	15-MAR-2000; 2000DE-01013847.		
XX	06-APR-2000; 2000DE-01019058.		
PR	07-APR-2000; 2000DE-01019173.		
PR	30-JUN-2000; 2000DE-01012529.		
PR	01-SEP-2000; 2000DE-01043826.		
XX	(EPIC-) EPIGENOMICS AG.		
PA	Olek A, Piepenbrock C, Berlin K;		
PI	WPI; 2001-602752/68.		
XX	The invention relates to a nucleic acid comprising a sequence of 18		
XX	bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with		
CC	bisulphite, of genes associated with tumour suppression and oncogenes		
CC	having a sequence taken from 536 (actually 533 since numbers 408, 458 and		
CC	500 are missing from the sequence listing) sequences (Ss) and sequences		
CC	complementary to (Ss). The nucleic acid may be a peptide nucleic acid-		
CC	oligomer (PNA) of at least 9 nucleotides and may form part of a set of		
CC	probes for detecting the cytosine methylation state and/or single		
CC	nucleotide polymorphisms and also to be used in an array for analysing		
CC	diseases associated with CpG dinucleotides e.g. cancers and tumours. The		
CC	probes can also be used in a method for ascertaining genetic and/or		
CC	epigenetic parameters for the diagnosis and/or therapy of existing		
CC	diseases or the predisposition to specific diseases, by analysing		
CC	cytosine methylations. The parameters may be compared to another set of		
CC	genetic and/or epigenetic parameters, the differences serving as basis		

CC for diagnosis and/or prognosis events which are disadvantageous to
CC patients. The present sequence is one of the 533 genomic sequences
CC derived from tumour suppressor genes and oncogenes. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC http://wipo.int/pub/published_pct_sequences

SQ Sequence 6292 BP; 1736 A; 99 C; 1304 G; 3153 T; 0 U; 0 Other;

Query Match	10.2%;	Score 151.2;	DB 4;	Length 6292;
-------------	--------	--------------	-------	--------------

Matches 619; Conservative 0; Mismatches 708

Matches 619; Conservative 0; Mismatches 708; Indels 13; Gaps 3;

[illegible]

QY	992	TAAAAACGAAACAAAAAGAAAGATAAATCTGCTGCATACAAAATCATACATTTAGACGACG	991
Db	3648	CAAAAAAAAAAAAAAAAAAAAAAAGATTAATAAAAAACAATTTTAAATPAAACGAAAAAAAAAAAAA	3589
QY	992	TAAATATTTCTGATGTTAAATGATTTTCAAAATAAAGTAAGTATGAGAGTAAATATAGTCT-	1050
Db	3588	AAAAAATATACGACACACAAAAATTAABAAAAAACTTTAAABAAAAAAATTAABAAAAAATATTT	3529
QY	1051	----GAATATGACGATTCATTAATATAGTGAAGAAAGAAAGATGATGAAACTTATGCGAATTT	1106
Db	3528	TTAAAAAAATTAATAAAAAACAAAAAAATTAATTAATAACGAAAAACGAAAAAAATATACC	3469
QY	1107	TAAAGCTATTTGCGAATATGACAAATTTCCAAAGATGAGAAAAACATAGAAATTTTAAGA	1166
Db	3468	AAAAAAAAAAAAAAAAAACGCAATTAABAAACAAATTTTAAATTAATTAATTAABAAAAACGA	3409
QY	1167	ACTAGAAAGATTGTATGAGAAAAAATGAATAATTTAGATGATTTAGATGAAAGAAATAGAAA	1226
Db	3408	AATTAACGATTTAATTAATAAAAACTTTACAAAAAAACAAATTAABAAAAAATTAATTAATAA	3349
QY	1227	ATCATCGAACAATTAATCTGAGAAAAATTAABAAAAAGAAAGAAATATGTAAAAAACAA	1286
Db	3348	AAAAATTTTAAATTAATTAABAAAAAABAAAAAATAGAAAAAABAAAAACCAAAACGATC	3289
QY	1287	GGATTAATTAATTTTAAACCAAAATGATAAAGTTTGTATGATGACGATTAATTAABAAATTA	1346
Db	3288	GAAAAAABAAATCTCTCATATTAATTTTTTAACTCTTCATTCGATTCGATCAAAAAA	3229
QY	1347	AAATGATAGCAGGTTAATA	1366
Db	3228	TAAATTAATTAATTAATTTCAAA	3209

RESULT 17

ID	standard; DNA; 6767 BP.
AAS46608	

AC AAS46608;

DT 18-DEC-2001 (first entry)

DE Tumour suppressor gene derived chemically modified sequence #3330.

Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;

cytosine methylation; ds.

Homo sapiens.

PN WO200168912-A2

AA PD 20-SEP-2001.

AA 15-MAR-2001
PF

AA
PR 15-MAR-2000; 2000DE-01013847.

PR	07-APR-2000; 2000DE-01019173.
PR	08-APR-2000; 2000DE-01019038.

PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
PR 01-SEP-2000; 2000DE-01043826.

PA (EPIG-) EPIGENOMICS AG.

XX	Olek A.	Piedenbrock C.	Berl
PI			

XX
DB
WPT: 2001-602752/68

XX	PT	Eracments of chemical}
----	----	------------------------

PT genes and oncogenes, useful in diseases associated with cytosine

XX
Page 1. See to no

XX

The invention relates to a nucleic acid comprising a sequence of 18 bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 438 and 500 are missing from the sequence listing) sequences (S8) and sequences complementary to (S8). The nucleic acid may be a peptide nucleic acid-oligonucleotide (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms and also to be used in an array for analysing diseases associated with CpG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases, by analysing cytosine methylations. The parameters may be compared to another set of genetic and/or epigenetic parameters, the differences serving as basis for diagnosis and/or prognosis events which are disadvantageous to patients. The present sequence is one of the 533 genomic sequences derived from tumour suppressor genes and oncogenes. Sequences with even numbered Seq ID numbers are the complementary sequence of the corresponding odd numbered sequence (e.g. ID 2 and ID1, ID 536 and ID 533, except for those whose partner sequence is missing). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pcc_sequences

Sequence 6767 BP; 1446 A; 207 C; 1601 G; 3513 T; 0 U; 0 Other;

Query Match	10.1%	Score 149.2	DB 4	Length 6767
Best Local Similarity	46.1%	Pred. No. 1.2e-12		
Matches 643, Conservative	0	Mismatches 743	Indels 8	Gaps 4

Qy	2	AAGAACCAACAAAGCGATCTAGAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAC	61
Db	3628	AATAAATATAAACCAACCCAAAABAAACCCAAAATCCGAAAABAAATTAACCAAAAAA	CT 356
Qy	62	AAAGCGATTTAGAACAGATAGACTTGGCTTAAAGAAAAGTTACAGAGACCAAGCGATT	121
Db	3568	AAA--CAATACACTCCAAAATTCACATAAAAAATTAATAATCAAAAAACAAATATCTCCAA	351
Qy	122	TAGAACACAGAGAGACTTGGCTTAAAGAAAAGTTGCAAGAACAAAGCGATCTTAGAACAG	181
Db	3510	AATATATATTAATTAATACTTAATCTAAACCTAACCTTCTTAAAAATATCTAATATTA	345
Qy	182	AGAGACGTGCTTAAGAAAAAGTTGCAAGAACAAACGATTTAGACAAAGAGACGTG	241
Db	3450	AAACCCGAAAAAATCTTATATATATAAAAAATCAAAAAAATATCTAAAAA	339
Qy	242	CTAAAGAAAAGTTGCAAGAACAAAGCGATTTAGAACAAATATGACTTGGTAAAGAA	301
Db	3390	CAAAAAAATCTAAAAAATAAAAAATAAAAAATCAAAAAAATCAAAAAATCAAAAAA	333
Qy	302	AGTTACAAGACGCAAGCGATTTAGAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAG	361
Db	3330	ACAAAAAATATCAAAAAACAAAAAATCAAAAAAATCAAAAAAATTAATAATCAAAA	327
Qy	362	AACACAAAGCGATTTAGAACAAAGAGACGCTGCTAAAGAAAAGTTGCAAGAACAA	421
Db	3270	ACATACAAAAACAAAAAATCAACAAAAATACATAAAAAATCAAAAAATCAAAAAATTA	321
Qy	422	GCGATTTAGAACAAAGAGACTTGGCTTAAAGAAAAGTTGCAAGAACAAACGATTTAG	481
Db	3211	AAAAAATAATACGAATCCAAAAATTAATAAAAAATCAAAAAACAAAAAATCAACAC	315
Qy	482	AAACAAGAGACGCTCTAATAGAAAAGTTGCAAGAACAAACGATTTTAGAACAGAC	541
Db	3151	AAACTTAACCAATATAAAAACCAAAAAACAAAAAATCAAAAAAATTAACAA	309
Qy	542	GACGTGCTAAAGAAAAGTTGCAAGAACAAACGCGATTTAGAACAAAGAGACGCTGA	601
Db	3091	AACAAAAATTAACAAAAACCAAAAAACAAAAAATTAACAAAAACCAAAAAACA	303
Qy	602	AAGAAAAGTTGCAAGACCAAAAGATTTTAGAACAAAGAAAGCTGATTCAGAAAAA	661

Db		3091	AAAAAAAAAATTCAAAAACAAAACCAAAACCTAACAAAACCAAAAAATAAAAAA	2972
Oy		662	ATTGAGAAAGAAAAAGAACATGAGATATATTAGCAGAGATTATATGTGGTTAG	721
Db		2971	TCAAAAAACAIAAAAAAAAAAACAACAAACCAAACCTAATATAAAAAACCAAAAA	2912
Oy		722	AAATACAGCATATGAACCTTCCATCAGAAAAATGAACTGGATATTATATCACATCAAT	781
Db		2911	AAACAAAAACAAAAAATAACAAAAACAAAAATTAACAAAACCAAAAAACAAAAAAT	2852
Oy		782	CTTCTTTACTCAGACACAGAAGGATAGTAGATTCCAGAGAAATATCTTATAT--	839
Db		2851	ATCAAAAAACAAAAAACAAACACAAAAATTAACAAAACCAAAAAATAAAAAATCAAAAC	2792
Oy		840	-AGAAAAAACCAATAGAAATCTATTACAACAATGTTGAAGACGAAGGATATACATA	898
Db		2791	AAAAAAAAAAAAACAAACCAATTAATTAIAAAAAACCAAAACCAAAAAACAAAAACAA	2732
Oy		899	AAGGACATCTTGAGAAAAGAAAGATGGTTCATTAACCAAGAACCAAAAAAGATTAAT	958
Db		2731	AACAAAAACATAAAAACAAAAATTAACAAACCAAAAAACAAATTCAAAAACAAAAACAAA	2672
Oy		959	CTGCTGACATTCACAAATCATATCATTAGACACATTAATTTCTGATGTTAATGATTTTC	1018
Db		2671	AAAAACCAACACAAAAACAAAAAATAATCACAACCAAAAAACGACAAAAATACA	2612
Oy		1019	AAATAGTAGTATGAGAGTAGAAATTAAGTGTGAATATGACGATTCTTAATAGATGAG	1078
Db		2611	AAACCAAAAAAIAAAAAAAAAAATAACAAACCAAAAAACATCAAAAAACAAAAATAAAA	2552
Oy		1079	AAGAAGATGATGAAGCTTAGACAAATTTAAGCTATTGTGCATATGACAAITTCAG	1138
Db		2551	ACAAAAATTAACAAACCAAAAAACAAATTAACAAAAACAAAAATACGAACCAAAAAATAAA	2492
Oy		1139	ATGAAGAAAAATCATGAATTTATTAAGAACTAGAAAGATTGATGAGAAAAATGAAATT	1198
Db		2491	AAAAAATAATCAAAAAACAAAAAATAATTAATTAACAAAAATTAACAAACCAAACTAAAA	2432
Oy		1199	TAGATGATTAGATGAAGAAATGAAAAATCATCAGAGATTAATCTGAAGAAAAATAA	1258
Db		2431	AAAAAATCAAAAAACAAAAAATAAAAAACAAATTAIAAAAAACCAATTAATTAACAA	2372
Oy		1259	AAAAAGAAAGAAATATGAIAAAAAACAAAGATTAATTTAAACCAATGATTAAGTT	1318
Db		2371	ACCAA--AAACAAAAAIAAAAAAAAAATCAAAAAAATTCMAATPAIAAACAAAAATTAACAAC	2314
Oy		1319	TGTATGATGACATATTAIAAAAAATTAIAAAATGTTAAGCAGTTAATPAAGAAAAAGAA	1378
Db		2313	AAAAACAAATTTAAACAAAAAATAACAAAAACAAATTAACAAACCAAAAAACAAAAA	2254
Oy		1379	AATTCATTAATCA 1392	
Db		2253	AAAAATTAIAACA 2240	
RESULT 18				
ABL05274/C				
ID	ABL05274	standard;	CDNA; 3662 BP.	
XX	ABL05274;			
AC				
XX				
DT	26-MAR-2002	(first entry)		
XX				
DE	Drosophila melanogaster expressed polynucleotide seq ID NO 10304.			
XX				
KM	Drosophila; developmental biology; cell signalling; insecticide;			
XX	pharmaceutical; gene; ss.			
OS	Drosophila melanogaster.			
PW				
XX	MO200171042-A2.			
PD	27-SEP-2001.			

```

XX 23-MAR-2001; 2001WO-US0092231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX MPI: 2001-656860/75.
DR P-PSDB: ABB61171.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 10304; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signaling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB872022). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 3662 BP; 661 A; 941 C; 421 G; 1639 T; 0 U; 0 Other;
SQ
Query Match 9.9%; Score 146; DB 4; Length 3662;
Best Local Similarity 46.7%; Pred. No. 3.5e-12;
Matches 603; Conservative 0; Mismatches 675; Indels 12; Gaps 4;
QY 24 ACAAGAGACGCTGCTAAAGAAAAGTTGCAGAACACAAACGATTTAGAACAAAGATAG 83
DB 1859 AAAAAAGAGATGAAATTAACAAGATCAAGGAAACCAAGATTAATGAAAAGGAAAG 1800
QY 84 ACTTGCTAAAGAAAAGTTTCAAGAGACGAAAGCCGATTTTGAACAAGAGACTTGCTTA 143
DB 1799 A--ACTTAAAGAGAGAGAAAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1742
QY 144 AGAAAAGTTGCAAGAACAAACAAAGCGATCTAGAACAGAGAGACGCTGCTAAAGAAAGTT 203
DB 1741 AGAGAAAGAGAAAGAAAAGAGAGAGAAAGAAAAGAGAGAGAAAGAAAAGAGAA 1682
QY 204 GCAGAACCAACAAACGATTTTGAACACAGAGAGACGCTGCTAAAGAAAGAAAGCA 263
DB 1681 GAGGAAAGAAAAGAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1626
QY 264 ACNAAAGCGATTTTGAACAAGATAGACTTCTTAAAGAAAAGTTTGAAGACGACCAAGCGA 323
DB 1635 GGAAAAGAGAAAGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAG 1571
QY 324 TTTAGAACCAAGAGAGACGCTGCTAAAGAAAAGTTGCAGAACCAAGCGATTTGAAGCA 383
DB 1570 GAAAAGAAAAGAGAGAGAAAGAAAGAGAAAGAAAGAGAAAGAAAGAGAAAGAG 1511
QY 384 AGAGAGACGCTGCTAAAGAAAAGTTGCAAGAACCAACGATTTTGAACAAGAGAGACT 443
DB 1510 AGAAAAGAAAGAGAAAGAAAGAAAGAGAAAGAAAGAGAAAGAAAGAGAGAG 1451
QY 444 TGCTTAAAGAAAAGTTGCAAGAACCAACGATTTTGAACAAGAGAGACGCTGCTAAAGCA 503
DB 1450 GAAAGAGAAAGAGAGAGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAG 1391
QY 504 AAAAGTTGCAAGAACCAACGATTTTGAACAAGAGAGACGCTGCTAAAGAAAAGTTGCA 563
DB 1390 AGAGAAAGAGAAAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAGAGAGAG 1331
QY 564 AGAACCAACCAAGCATTGTAACAAGAGAGACGCTGCTAAAGAAAAGTTGCAAGAGACAGA 623

```

```

DB 1330 GAAAGAAAAGAGAGAAAGAGAAAGCAGAGAGAAAGAAAGAAAGAGAAAGAGAGA 1271
QY 624 AAGAGATTTTGAACAAGAAAGCGCTGATACCAAAAAAATTTAGAAAAGAAAAGAGACA 683
DB 1270 GAAAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1212
QY 684 TGAGATATATTTAGACAGAGATTTATATATGCTGCTTGAAGAAATACAGCTATAGACTTCC 743
DB 1211 AGAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1152
QY 744 ATCAAAAAATGACGTGATATTTATATACATCAATCTTTCTTACCTCAGACACAG 803
DB 1151 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1092
QY 804 AGGGAATGTGAGATTCACAGAAATATCTTATATAGAAAACAAATAGAAATCTAT 863
DB 1091 AGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1032
QY 864 TACAACAAATTTGAAAGAGAGAGAGATATATATTAAGAGACATCTTGAAGAAAAGAGAG 923
DB 1031 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
QY 924 TGCTTCAATTAACCCAGAACCAAAAGAGATTAATCTGCTGACATACAAATCATACATT 983
DB 971 AGAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 912
QY 984 AGAGACGCTAAATTTTCTGATGTTATGATGATTTTCAATTAAGTATGAGATGAAAT 1043
DB 911 ACAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852
QY 1044 AAGTCTGAATATGACGATTCATTAATGATGAGAGAGAGAGATGAGAGACTTGAAGCA 1103
DB 851 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
QY 1104 ATTTAAGCTTATTTGTGCAATATGACATTTTCCAGATGAGAGAGAGAGAGAGAG 1163
DB 791 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732
QY 1164 AGAACTAGAAATTTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223
DB 731 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672
QY 1224 AAAATCATCAGAGAGATTAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1283
DB 671 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 612
QY 1284 AAAGATATATATTTTAAACCAATGATTA 1313
DB 611 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 582

```

RESULT 19
 AAS45347/C
 ID AAS45347 standard; DNA; 9539 BP.
 XX
 AC AAS45347;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Chemically pretreated complementary DNA associated with cell cycle #26.
 XX
 KW Cell cycle; human; Cpg dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Levy body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.
 XX
 OS Homo sapiens.
 XX
 XX WO200168911-A2.
 PN

[illegible]

XX DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
 KW PNA: cytosine methylation state; SNP: retroviral infection; gene; ds;
 KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
 KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
 KW immunological disorder; Werner syndrome; developmental disorder;
 KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
 KW neurodegenerative disorder; Maardenburg syndrome; Niemann-Pick disease;
 KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
 KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;
 KW polyglutamine disorder; solid tumour.
 XX Unidentified.
 OS WO200192565-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 06-APR-2001; 2001WO-EP003973.
 XX
 XX 06-APR-2000; 2000DE-01019058.
 XX 07-APR-2000; 2000DE-01019173.
 XX 30-JUN-2000; 2000DE-01032529.
 XX 01-SEP-2000; 2000DE-01043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-090046/12.
 XX
 XX New nucleic acids or oligomers, useful for diagnosing or treating
 PT diseases associated with DNA transcription, e.g. immunological disorders,
 PT Werner syndrome, psoriasis, myocardial infarction, solid tumors or
 PT cancer.
 XX
 XX Claim 1; SEQ ID NO 54; 32p; English.
 XX
 XX The invention relates to a nucleic acid, which comprises a segment of the
 CC chemically pretreated DNA of genes associated with DNA transcription from
 CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
 CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
 CC to the chemically pretreated DNA of genes associated with DNA
 CC transcription. The set of oligomer probes are useful for detecting the
 CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
 CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
 CC diagnosing or treating diseases associated with DNA transcription
 CC (particularly with the methylation status), e.g. adenosine deaminase
 CC deficiency, viral infection, retroviral infection, Sezary syndrome,
 CC haematological disorders, immunological disorders, Werner syndrome,
 CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
 CC neurological disorders, neurodegenerative disorders, Maardenburg
 CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
 CC infarction, hypertension, angiodysplasia, erythropoiesis, congenital heart
 CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
 CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
 CC associated genomic DNA molecules of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification but
 CC was obtained in electronic format directly from the printed patent
 CC Office
 CC
 XX
 SQ Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 U; 0 Other;

Query Match 9 4%; Score 139.4; DB 6; Length 9539;
 Beest Local Similarity 45.6%; Pred. No. 2.8e-11;
 Matches 607; Conservative 0; Mismatches 716; Indels 8; Gaps 3;

QY 72 AGAACAATGACCTGCTAAAGAAAGTTACAAGAGCAAGACGATTTAGAACAGA 131
 Db 1343 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1284
 QY 132 GAGACTTGCTAAAGAAAGTTGCAAGAACAAAGCATCTAGAACAGAGAGAGCTGC 191

Db 1283 AAAATCAAAACCCGAAACCATTAATTTCTAAAAAATTTAAAAACAAATATACACAA 1224
 QY 192 TAAAGAAAGTGGAGAGCAACAAAGCATTTAGAACAGAGAGAGCTGTTAAAGAAA 251
 Db 1223 TAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1164
 QY 252 GTTCCAGAAACAAAGAGCGATTTAGAACAAATAGACTGCTTAAAGAAAGTTACAGA 311
 Db 1163 AAAACAAAAAACTTAATTC-CTACATACATCCACCTCCGAAATCTCTTAAAAAA 1105
 QY 312 GCAGCAAGCGATTTAGAACAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAAG 371
 Db 1104 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1045
 QY 372 CGATTTGAACAGAGAGAGCTGCTTAAAGAAAGTTGCAAGAACAAAGCGATTTGA 431
 Db 1044 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 985
 QY 432 ACAAGAGAGACTTGCTTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAG 491
 Db 984 AAAAAAAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 925
 QY 492 AGTGCTTAAAGAAAGTTGCAAGAACAAAGCGATTTAGAACAGAGAGAGAGCTGCTAA 551
 Db 924 AAAAAAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 865
 QY 552 AGAAAGTGGCAAGAACAAAGCGATTTAGAACAGAGAGAGAGAGCTGCTTAAAGAAAGT 611
 Db 864 AAATCAGCTTAAACCAAAATTTCAAAACCAAGCTTAAACCAAAATCAGCTTAAAGT 805
 QY 612 GCAAGAGAGCAAGAAAGATTTAGAACAAAGAGAGCTGATACGAAAAAATTTAGAAAG 671
 Db 804 ACAAAAAAAGAAAAAAGTAACTAACAAAGCTTAAACAGACCTTAATTAATTAATTAATTA 745
 QY 672 AAAAAAGAAATGAGATTAATTAAGAGAGATTTAATGAGCTGTTAGAAATCAAGC 731
 Db 744 AAAAAAGTAAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 691
 QY 732 TATGAAGCTTCCATCAAGAAATGAAGTGAATATTAATTAATTAATTAATTAATTA 791
 Db 690 AATGTACACTTACATCTCAAGCTTAAACAAAGAAAGCTTAACTCAAAAAAATTA 631
 QY 792 TCAGAGACAGAGAGAGATTTAGAGATTTCCAGAGAAATATCTTAATTAAGAAAAACAA 851
 Db 630 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 571
 QY 852 TAGGAATCTATTACAACAAATGTTGAAGAGCGAAGGATTAATTAAGAGAGATCTTGA 911
 Db 570 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 511
 QY 912 AGAAAGAAAGATGCTTCAATTAACCAAGAACAAAGAAAGATTAATCTGCTGACATCA 971
 Db 510 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 451
 QY 972 AAATCATATATTAAGACAGATTAATTTCTGATGTTAATGATTTCCAAATTAAGTAA 1031
 Db 450 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 392
 QY 1032 TGAGATGAATTAAGTCTGATTAATGACGATTCATTAATGATGAAGAGAGATGATGA 1091
 Db 391 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 332
 QY 1092 AGACTTGAAGAAATTTAAGCTTATGTCATATGACATTAATTTCCAGATGAAGAAACAT 1151
 Db 331 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 272
 QY 1152 AGAAATTTAATTAAGAACTAGAGATTTGATGAGAAATTAATTTAGATGATTTGA 1211
 Db 271 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 212
 QY 1212 TGAAGAAATGAAGAAATCATCAGAGAAATTAATGAGAAAGAAAAAAGAGAAAGAA 1271
 Db 211 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 152

QY	1272	ATATGAAAAAACAAGGATTAATTAATTTAAACCAATGATTAAGTTGTATGATGACA	1331
Db	151	AAAAAAAAAAAAAAAAATATTAATAAAAAAAAAAAAAAAAAATTAATAATTAATTAATAAT	92
QY	1332	TATTAATAATAATTAATAATAATGATTAAGCAGGTTAATTAAGGAAAAAGGAAAAATTCATTAATC	1391
Db	91	AAATTTTC	32
QY	1392	ATTGTTTCATA	1402
Db	31	TATATCTAAAA	21
RESULT 21			
ID	ABLJ33697/C	ABLJ33697 standard, DNA, 6668 BP.	
XX	ABLJ33697;		
AC			
DT	26-MAR-2002	(first entry)	
XX			
DE	Human immune system associated gene SEQ ID NO: 1670.		
XX			
KW	Human; immune system disease; cytosine methylation; antiaesthetic;		
KW	antiartherosclerotic; anlaemic; cytosolic; noctropic;		
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;		
KW	antirheumatic; antierthritic; antidiabetic; antiporiatic;		
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;		
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;		
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;		
de.			
XX			
OS	Homo sapiens.		
XX			
XX	WO200200928-A2.		
XX			
XX	03-JAN-2002.		
XX			
XX	02-JUL-2001; 2001WO-EP007537.		
XX			
PR	30-JUN-2000; 2000DE-01032529.		
PR	01-SEP-2000; 2000DE-01043826.		
XX			
XX	(EPIG-) EPIGENOMICS AG.		
XX			
PI	Olek A, Piepenbrock C, Berlin K;		
XX			
DR	WPI; 2002-130909/17.		
XX			
PT	Nucleic acid comprising fragment of chemically modified gene, useful for		
PT	diagnosis and treatment of diseases associated with abnormal cytosine		
PT	methylation.		
XX			
PS			
XX	Claim 1; SEQ ID NO 1670; 32pp + Sequence Listing; German.		
XX			
CC	The present invention provides a number of human immune system associated		
CC	genes which are modified by the methylation of cytosine. The sequences		
CC	can be used in the diagnosis and treatment of immune system disorders,		
CC	including eye diseases such as retinopathy, neovascular glaucoma and		
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid		
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,		
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel		
CC	diseases. The present sequence is a gene of the invention		
XX			
XX			
XX	Sequence 6668 BP; 1329 A; 328 C; 1736 G; 3274 T; 0 U; 1 Other;		
XX			
QY	Query Match	9.3%; Score 137.8; DB 6; Length 6668;	
QY	Best Local Similarity	44.8%; Pred. No. 4.8e-11;	
QY	Matches 621; Conservative	0; Mismatches 757; Indels 9; Gaps 2;	
QY	25	CAAGAGACGCTGCTTAAGAAAGTTCGACACACAAACCGATTGACAAAGATGAC	84

[illegible]

```
Oy 1164 AGAACTAGAGATTGTGATGAGAAAAATTTAGTGTGATTTAGTGAAGAAATGA 1223
Db 2721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2662
Oy 1224 AAATCATGAGAGATTATCTGAAAGAAAAATTAAGAAAGAAATATGAAAAAC 1283
Db 2661 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2602
Oy 1284 AAAGATATATATTTTAAACCAATGATAAAGTTGTATGATGACATATTAATAATA 1343
Db 2601 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAATCAAAAAAAAAAAAAAAAAAAC 2542
Oy 1344 TAAAAATGATAGCAGCTTAATTAAGAAAGAAAAATTCATTAATCATTTTCATAT 1403
Db 2541 AAATATCTATTTCTTAAACGTAACCTAATTCACATTAATAAAAAACCAATTTTAAATA 2482
Oy 1404 ATTTGAC 1410
Db 2481 AATAAAC 2475

RESULT 22
AAK33181
ID AAK33181 standard; DNA; 6644 BP.
AC AAK33181;
XX 25-JUN-1999 (first entry)
DT
XX Baee sequence of the plasmid pRx-ires-bar.
XX
XX Cowpox virus; bar; viral vector; expression; apoptosis; resistance; crmA;
KM autoimmune disease; graft rejection reaction; inflammation;
KM inflammatory disease; ss.
XX
XX Synthetic.
OS Cowpox virus.
XX
XX MO9913073-A2.
XX
XX 18-MAR-1999.
XX
XX 07-SEP-1998; 98MO-JP004010.
XX
XX 08-SEP-1997; 97JP-00259235.
XX
XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX
XX Hamada H;
XX
XX WPI; 1999-243728/20.
XX
XX New apoptosis-resistant virus-sensitive cell.
XX
XX Example 1; Page 38-41; 51p; English.
XX
XX The present invention describes an apoptosis-resistant virus-sensitive
XX cell line into which an apoptosis resistance gene has been introduced.
XX The recombinant viruses generated are capable of expressing apoptosis-
XX associated genes. These can then be used in a variety of diseases for
XX which the induction of apoptosis by gene transfer, or where the
XX inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
XX are useful as vectors for gene therapy which can be applied to cancer
XX therapy for destroying cancer cells selectively, the treatment of
XX autoimmune diseases and graft rejection reaction, and apoptosis induction
XX therapy for inflammatory cells in inflammatory diseases. Prior arts have
XX encountered the problem where if an adenovirus vector capable of
XX expressing an apoptosis-associated gene is introduced into animal cells,
XX the cells producing the virus will be destroyed because the period of
XX time required to induce cell death by apoptosis is shorter than that
XX required to replicate and produce the virus, resulting in failure to
```

```
CC obtain a recombinant virus having the integrated apoptosis-associated
CC gene. In this invention an apoptosis-resistant 293 cell line (having an
CC apoptosis resistant gene introduced) is established and overcomes the
CC problem. The present sequence represents the base sequence of the plasmid
CC pRx-ires-bar, which contains the cowpox virus bar gene, and is used in an
CC example from the present invention
XX
SQ Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 U; 0 Other;
Query Match 9.2%; Score 136.4; DB 2; Length 6644;
Best Local Similarity 49.6%; Pred. No. 7.7e-11;
Matches 350; Conservative 0; Mismatches 356; Indels 0; Gaps 0;
Oy 1 CAAGACACCAACGATCTCTGAACAGAGAGCGTGTAAAGAAAAGTTCAGACACAA 60
Db 3738 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3797
Oy 61 CAAAGCATTTAGAACAGATAGACTTGCTAAAGAAAAGTTACAAAGACCAAGCGAT 120
Db 3798 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3857
Oy 121 TTAGACACAGAGACTTGCTTAAAGAAAAGTTGCAAGACCAAGCGATCTAGACAA 180
Db 3858 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3917
Oy 181 GAGAGACGTGCTAAAGAAAAGTTGCAAGACCAAGCGATTTAGAACAAAGAGACGT 240
Db 3918 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3977
Oy 241 GCTAAAGAAAAGTTGCAAGACCAAGCGATTTAGAACAGATAGACTTGCTAAAGAA 300
Db 3978 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4037
Oy 301 AAGTTACAGAGCGCAAGCGATTTGACACAGAGACGTGCTAAAGAAAAGTTGCA 360
Db 4038 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4097
Oy 361 GAACACAAAGCGATTTAGAACAGAGAGCGTGTAAAGAAAAGTTGCAAGACCAAG 420
Db 4098 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4157
Oy 421 AGCGATTTAGAACAGAGAGACTTGCTAAAGAAAAGTTGCAAGACCAAGCGATTTA 480
Db 4158 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4217
Oy 481 GAACACAGAGAGCGTGTAAAGAAAAGTTGCAAGACCAAGCGATTTGACACAGAG 540
Db 4218 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4277
Oy 541 AGACGTGCTAAAGAAAAGTTGCAAGACCAAGCGATTTAGAACAGAGAGCGTGT 600
Db 4278 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4337
Oy 601 AAAGAAAAGTTGCAAGAGCAAGAGATTTAGAACAAAGAAAGCTGATACGAAAAA 660
Db 4338 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4397
Oy 661 AATTTAGAAAAGAAAAGCAATGAGATATATTAGACAGGATT 706
Db 4398 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAATTGTCGAAGCTT 4443

RESULT 23
AAK33182
ID AAK33182 standard; DNA; 7372 BP.
AC AAK33182;
XX 25-JUN-1999 (first entry)
DT
XX Baee sequence of the plasmid pRx-Bcl-xl-bar.
XX
XX Cowpox virus; bar; viral vector; expression; apoptosis; resistance; crmA;
```

KW bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer;
KW autoimmune disease; graft rejection reaction; inflammation;
KW inflammatory disease; ss.

OS Synthetic.

OS *Homo sapiens*.

PN WO9913073-A2.

PD 18-MAR-1999.

07-SEP-1998; 98WO-JP004010.

08-SEP-1997: 97JP-00259235

PA (BPRG-) BPR GENCELL, ASIA PACIFIC INC

XX
PT. Hamada H.

XX WPB - 1000-243720/20
DB

You complete registration at the registration desk

XX
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840

The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis-associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pPR-Bcl-xl-bar, which contains the human Bcl-xl gene, and is used in an example from the present invention.

Sequence 7372 BP; 2353 A: 1749 C: 1649 G: 1621 T: 0 U: 0 Other:

Query Match	9.2%	Score 136.4	DB 2	Length 7372
-------------	------	-------------	------	-------------

```

Matches 350; Conservative 0; Mismatches 356; Indels 0; Gaps 0;

```

[illegible]

QY	301	AAAGTTACAAGAGCGCAAGAGCGATTTTAGAACACAGAGAGAGCTGTTAAAGAAAGTTGCCAA	360
Db	4766	AA	4825
QY	361	GAACAACAAGCGATTTTAGAACAGAGAGAGCGTGCTAAAGAAAGTTGCCAAGAACACAA	420
Db	4826	AA	4885
QY	421	AGCGATTTAGAACAGAGAGAGCTTGCTAAAGAAAGTTGCCAAGAACACAAAGCGATTTA	480
Db	4886	AA	4945
QY	481	GAACAACAAGAGAGCGTGCTAAAGAAAGTTGCCAAGAACACAAAGCGATTTAGAACAAAG	540
Db	4946	AA	5005
QY	541	AGAGCTGCTAAGAAAGTTGCCAAGAACACAAAGCGATTTAGAACAGAGAGAGCGTGCT	600
Db	5006	AA	5065
QY	601	AAAGAAAGTTGCCAAGAGCAGCAAGAGATTTAGAACAAAGAGAGCGTGATTACGAAAAAA	660
Db	5066	AA	5125
QY	661	AATTTAGAAAGAAAAAGAAACATGAGATATATTTAGCAAGAGATT	706
Db	5126	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATGGCTCCAGGCTT	5171

RESULT 24
AAX33180

AC AAX33180;

DT 25-JUN-1999 (first entry)

DE Cowpox virus bsr full length gene sequence.

Cowpox virus; bsr; viral vector; expression; apoptosis; resistance; crmA;

KW autoimmune disease: graft rejection reaction: inflammation:

XX inflammatory disease; ss.

Compox V1718.

PN WO9913073-A2.
YY

PD 18-MAR-1999.

PF 07-SEP-1998;

PR 08-SEP-1997; 97JP-00259235.

PA (RPRG-) RPR GENCELL ASIA PACIFIC INC.

PI Hamada H;

DR WPI; 1999-243728/20.

PT New apoptosis-resistant virus-sensitive cell.

Example 1: Page 34-38; 51pp; English.

The present invention describes an ar

CC The present invention describes an apoptosis-resistant virus-sensitive
CC cell line into which an apoptosis resistance gene has been introduced.
CC The recombinant viruses generated are capable of expressing apoptosis-
CC associated genes. These can then be used in a variety of diseases for
CC which the induction of apoptosis by gene transfer, or where the
CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
CC are useful as vectors for gene therapy which can be applied to cancer
CC therapy for destroying cancer cells selectively, the treatment of
CC autoimmune diseases and graft rejection reaction, and apoptosis induction

CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
CC encountered the problem where if an adenovirus vector capable of
CC expressing an apoptosis-associated gene is introduced into animal cells,
CC the cells producing the virus will be destroyed because the period of
CC time required to induce cell death by apoptosis is shorter than that
CC required to replicate and produce the virus, resulting in failure to
CC obtain a recombinant virus having the integrated apoptosis-associated
CC gene. In this invention an apoptosis-resistant 23 cell line (having an
CC apoptosis resistant gene introduced) is established and overcomes the
CC problem. The present sequence represents the comox virus bsr gene which
CC is used in an example from the present invention

SQ Sequence 7797 BP; 2542 A; 1760 C; 1656 G; 1839 T; 0 U; 0 Other;

Query Match	9.28;	Score 136.4;	DB 2;	Length 7797;
-------------	-------	--------------	-------	--------------

Best Local Similarity 49.6%; Pred. No. 7.6e-11;

```
Matches 350; Conservative 0; Mismatches 356; Indels 0; Gaps 0;
```

[illegible]

RESULT 25
AAx33184
ID AAx33184 standard; DNA; 7996 BP.

AC
XX
AA33184;

DT 25-JUN-1999 (first entry)

DE Base sequence of the plasmid pRx-Bcl 2-1-hcd 25.

KM Cowpox virus; bsr; viral vector; expression; apoptosis; resistance; crmA;

KW bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer

KM autoimmune disease; graft rejection reaction; inflammation;
KM inflammatory disease; os

KW inflammatory disease; vv.
XX

50 yy

Homo sapiens.

XX 8

PN WO9913073-A2.

3000

PD 18-MAR-1999
YY

XX 07-SEP-1998. 98W0-TD004010
PF

PF 07-SEP-1998; 98WC-JP004010.
XX

PR 08-SEP-19

XX : 8
C : 1
F : 1
C : 1
C : 1
C : 1
C : 1

PA (RPRG-) RPR GENCELL ASIA PACIFIC INC.

2000

XX
XX
DR WPI; 1999-243728/20.
XX
PT
XX New apoptosis-resistant virus-sensitive cell.
PS Example 3; Page 46-49; 51pp; English.
XX
XX The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis-associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRX-Bcl 2-1-hcd 25, which contains the human Bcl-2 gene, and is used in an example from the present invention

SQ Sequence 7996 BP; 2463 A; 2015 C; 1829 G; 1689 T; 0 U; 0 Other;

Query Match	9.28;	Score 136.4;	DB 2;	Length 7996;
-------------	-------	--------------	-------	--------------

Best Local Similarity 49.6%; Pred. No. 7.6e-11;

Matches 350; Conservative 0; Mismatches 356; Indels 0; Gaps 0;

[illegible]

DT 12-JUL-2002 (first entry)
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 2180.
 DE
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 XX 01-SEP-2001; 2001WO-EP010074.
 XX
 XX 01-SEP-2000; 2000DE-01043826.
 XX
 XX 05-SEP-2000; 2000DE-01044543.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 XX
 PA Olek A, Piepenbrock C, Berlin K, Guetig D;
 PI WPI; 2002-371829/40.
 DR
 XX
 XX Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.
 XX
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention
 CC
 XX
 XX
 SQ Sequence 778 BP; 614 A; 29 C; 24 G; 111 T; 0 U; 0 Other;
 Query Match 9.0%; Score 132.8; DB 6; Length 778;
 Best Local Similarity 48.5%; Pred. No. 2.9e-10;
 Matches 365; Conservative 0; Mismatches 387; Indels 0; Gaps 0;

DB 166 ATTACGAAAAAAAAAATTATAAAAAAAAAAAAAAAAAATTACG 245
 QY 878 AAGCGAAGGAGATTTACCTAAGACATCTTGAAGAAAGAGTGGTCAATTAAC 937
 DB 246 AAAAAAAAAAATTTATAAAAAAAAAAAAAAAAAATTTATAAAAAAAAAATTCGAAAAA 305
 QY 938 CAGAACAAAGAGATTAATCTGCGCATTCGCAAAATCATCATTTAGAGACGTAAATA 997
 DB 306 AAAAAATTTACGAAAAAAAAAAAAAAAAAATTTACGAAAAAAAAAAAAAAAAAAT 365
 QY 998 TTTCTGATGTTAATGATTTTCAATAGTAACTAGAGATGAATTAAGTCTGATATG 1057
 DB 366 TTACAAAAACGAAAAATTTATAAAAAAAAAATTTACGAAAAAATTTACGAAAAA 425
 QY 1058 ACGATTCATTAATGATGAGAGAGATGATGAAGACTTACAGCAATTTAAAGCTATTG 1117
 DB 426 AAAAAAAAAATTCGAAAAAAAAAAAAAAAAAATTTATAAAAAAAAAAAAAAAAAAT 485
 QY 1118 TCGAATATGACAAATTTCCAGATGAGAGAAACATAGCAATTTATTAAGAACTAGAAAGATT 1177
 DB 486 TACGAAAAAAAAAAAAAAAAAATTTATAAAAAAAAAATTTACGAAAAAATTTACGAAAAA 545
 QY 1178 TGATAGAGAAAAATGAAATTTAGATGATTTAGATGAGAAATAGAAAAATCATCAGAAG 1237
 DB 546 AAAAAATTTATAAAAAAAAAAAAAAAAAATTTACGAAAAAATTTACGAAAAAATTTACGAAAAA 605
 QY 1238 AATTATCTGAAGAAAAATTAAGAAAAAGAAAGAAATATGAAAAAACAAGATAATTAATT 1297
 DB 606 AATTATCGAAAAAATTAAGAAAAAATTTACGAAAAAATTTACGAAAAAATTTACGAAAAA 665
 QY 1298 TTAACCAAAATGATAAAGTTTGTATGATGACATTTAAAAAATTTAAAAATGATTAAC 1357
 DB 666 AAAAAAAAAAATTTACGAAAAAATTTATAAAAAAAAAATTTATTAATA 725
 QY 1358 AGCTAATAGAGAAAAAGAAAAATTCATTAATA 1389
 DB 726 AAAAAAAAAATTTATAAAAAAAAAATTTACAA 757

RESULT 29
 ADS89401/C
 ID ADS89401 standard; DNA; 5493 BP.
 AC ADS89401;
 XX
 XX 18-NOV-2004 (first entry)
 DT
 XX
 XX Oligonucleotide of the invention SEQ ID NO:417.
 DE
 XX
 KW ss; cell proliferative disorder; breast; methylation; cytostatic;
 KW gene therapy; single nucleotide polymorphism; SNP.
 XX
 XX Unidentified.
 OS
 XX
 PN WO2004035803-A2.
 XX
 PD 29-APR-2004.
 XX
 PF 01-OCT-2003; 2003WO-EP010881.
 XX
 PR 01-OCT-2002; 2002DE-01045779.
 XX
 PR 07-JAN-2003; 2003DE-01000096.
 XX
 PR 17-APR-2003; 2003DE-01017955.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 XX
 PI Fockens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;
 PI Nimrich I, Kujan T, Schmitt A, Schmitt M, Look MP, Marx A;
 XX
 DR WPI; 2004-348468/32.
 XX
 PT Predicting responsiveness of a subject with breast cell proliferative
 PT disorder, useful for treating or differentiating breast cell

PT proliferative disorders comprises analyzing methylation pattern of a
PT genomic DNA from the subject.

XX Claim 25; SEQ ID NO 417; 104pp; English.

CC The invention relates to a novel method for predicting the responsiveness
CC of a subject with a cell proliferative disorder of the breast tissues to
CC a therapy comprising analysing the methylation pattern of a target
CC nucleic acid by contacting at least one of the target nucleic acids in a
CC biological sample obtained from the subject prior to or during treatment.
CC The method of the invention has cytosatic activity, and may have a use
CC in gene therapy. The set of oligonucleotides comprising at least two of
CC the oligomers are useful for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
CC method, nucleic acid, oligonucleotides, and kit are useful for the
CC treatment, characterisation, classification and/or differentiation, of
CC breast cell proliferative disorders. The method is also useful for
CC predicting the responsiveness of a subject with a cell proliferative
CC disorder of the breast tissues to a therapy. The present sequence is used
CC in the exemplification of the invention.

XX Sequence 5493 BP; 851 A; 174 C; 1692 G; 2776 T; 0 U; 0 Other;

Query Match 8.9%; Score 131.4; DB 13; Length 5493;
Best Local Similarity 46.3%; Pred. No. 4e-10;
Matches 503; Conservative 0; Mismatches 581; Indels 3; Gaps 2;

```

QY 306 ACAAGAGCGCAACGATTAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGA 365
DB 5475 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5416
QY 366 ACAAGCGATTAGAACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGA 425
DB 5415 AAAAAACGCAAAAAAAAAAAAAAAAAATCAACGAAAAACAAACGCGAATTAATAAAAAACAAA 5356
QY 426 TTTAGACAGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAAC 485
DB 5355 ACGATCGAAATATATTAACAAAAAAGTTGCAAGAACAAAGCGAATTAATAAAAA 5296
QY 486 AGAGAGAGCGTCTAAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGCG 545
DB 5295 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5236
QY 546 TGCTTAAGAAAGTTGCAAGAACAAAGCGATTAGAACAGAGAGAGCGTCTAAAG 605
DB 5235 AAAAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAA 5176
QY 606 AAAGTTGCAAGAGCAAGAGATTAGAACAAAGAGCGTCTAGTACGAAAAAAATTT 665
DB 5175 AAAAAACAAAAAACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 5116
QY 666 AGAAAGAAAGAAAGCAATGAGATATATAGACAGATTTATATGCTGTTTGAAT 725
DB 5115 ACGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5056
QY 726 ACCAGCTATAGAACTTCATCGAAGAAATGAACGTGATATTTATACCAATCAATCTTC 785
DB 5055 AAAAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4996
QY 786 TTTACCTGAGCAACAGAGGGAATAGTAGAGATTCAAGAAATATCTATATAA -GAA 843
DB 4995 AATACAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4936
QY 844 AAAAAAATTAAGAAATCTATTAACAATAATGTTGAAGACGAAGGATATCATTAAGA 903
DB 4935 AAAAAATTAATCTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4876
QY 904 CATCTTGAAGAAAGAAAGATGTTCAATTAACAGAAACAAAGAAAGATTAATCTGCT 963
DB 4875 CAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4816
QY 964 GACATACAAATCATATAGAGAGAGTAAATATTTCTGATGTTATGATTTTCAATA 1023

```

```

DB 4815 GAAATTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4756
QY 1024 AGTAAATAGAGATGAATTAAGTGC-TGAATAGAGATTCATTATAGATGAAGAGA 1082
DB 4755 TACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4696
QY 1083 AGATGATGAAGACTTAGACGATTAAGCTATGTCATATGATGATTAAGATTAAGATGA 1142
DB 4695 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4636
QY 1143 AGAAACATAGGAATTTATTAAGAACTAGAAAGATTGATAGAGAAAAATGAATAATTGA 1202
DB 4635 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4576
QY 1203 TGATTTGATGAAGATTAAGAAATATCATCGAAGAAATTAATCTGAAGAAAAATTAATAA 1262
DB 4575 ACGAAACAAACCAAAATCGAAACCCCAACAAAAAATATCCAAATACAAATCAATAAAT 4516
QY 1263 AGAAAGAAATATGAAAAAATAAAGATTAATTTTAAACCAATGATTAAGATTTGTA 1322
DB 4515 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4456
QY 1323 TGATGAGCATTTTAAATAATTAATAATGATAGCGAGTTAATAGAAAGAAAAATTT 1382
DB 4455 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4396
QY 1383 CATAAA 1389
DB 4395 CCCCAA 4389

RESULT 30
ID AAS46686 standard; DNA; 7442 BP.
XX AAS46686;
AC 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #409.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytosatic; cancer;
KW cytosine methylation; de.
XX
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PE 15-MAR-2001; 2001WO-EP002955.
XX
PR 15-MAR-2000; 2000DE-01013847.
PR 06-APR-2000; 2000DE-01019058.
PR 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIDENOMICS AG.
XX
PI Olek A, Pienbrock C, Berlin K;
XX
DR WPI; 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumor suppressor
PT genes and oncogenes, useful in designing primers and probes for analyzing
PT diseases associated with cytosine methylation state e.g. cancer.
XX
PS Claim 1; SEQ ID NO 409; 277p; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with

```

Query Match	8.9%	Score	131.4	DB	4	Length	7442
Best Local Similarity	46.7%	Pred	No. 3.9e-10				
Matches	565	Conservative	0	Mismatches	626	Indels	18
						Gaps	4

Oy	914	AAAAAA-----AGATGGTCAATAAAACGAGACAAAGAAAGATTAATCGC	962
Db	2359	AAAAAAAAATTAATTTTAAAAAAAAAAAAAAAAATCTAABAAAAAAAAATTAATCTTA	2300
Oy	963	TGACATACAAATCATACATTAGAGACGTAAATTTTCTGATGTAATGATTTCAAT	1022
Db	2299	AAATTAATTAATTTTAAAAAAAAAAAAAAAAATTAATCTAAAAAAAAAAAACTTA	2240
Oy	1023	AAGTAAGATGAGGATGAAATTAAGCGCTGATATGACGATTCATTAAATAGATGAAGA	1082
Db	2239	AAAAAAAAATTAATCTAATAAAAAAAAAATTAATTAATCTAATAAAAAAAAAATTAATCTTA	2180
Oy	1083	AGATGATGAGACCTTAGACGAATTTAAGCCTATTGTGCATATGACAAATTTCCAGATGA	1142
Db	2179	AAAAAAAAATTAATTTTAAAAAAAAATTAATTAATTAATTTTAAAAAAAAA-ATAAATAATCTAA	2121
Oy	1143	AGAAATCTAGGAATTTTATTAAGACTTGAAGATTTGTATAGAAAAATGAAAAATTT--A	1200
Db	2120	AAAAAAAAATTAATCTAATAAAAAAAAAATTAATTAATTAATAAAAAAAAAATTAATCTTA	2061
Oy	1201	GATGATTTAGATGAAGATAGAAAAATCATCAGACGATTAATCTGAAGAAAAATTA	1260
Db	2060	AAAAAAAAAAAAATCTTAATAAAAAAAAAACAATCTAATAAAAAATCTAATAAAACCTAAT	2009
Oy	1261	AAAGAAAGAAATATGAAAAAACAAAGATTAATTTTAAACCAATGATTAAGTTTG	1320
Db	2000	AAAAAAACGCAACTCTAATAAAAAAAAA---AAAAAAATTAATAAAAAAAAAAAAAATCTTA	1949
Oy	1321	TATGATGACATATTAATAAAAAATTAATAATGATTAAGCAGTTTAATTAAGAAAAAGAAAA	1380
Db	1944	AAAAAAAAAATTTAATAAAAAATCTTAATAATAAAACCAATCTAATAAAAAAAAAAAAA	1885
Oy	1381	TTCTATAAA 1389	
Db	1884	CTCAAAAA 1876	

Search completed: December 20, 2005, 05:54:22
Job time : 1151.77 secs

THIS PAGE BLANK (USPTO)